

the CATO JOURNAL

AN INTERDISCIPLINARY JOURNAL OF PUBLIC POLICY ANALYSIS

ARTICLES

- | | |
|--|---|
| Anton D. Lowenberg | A Post-Apartheid Constitution for South Africa:
Lessons from Public Choice |
| Peter J. Hill | Environmental Problems under Socialism |
| Marilyn R. Flowers | The Political Economy of Mandated Spending |
| Kenneth W. Costello
and David Haarmeyer | Reforming the Bonneville Power Administration |
| Gary M. Anderson and
Robert D. Tollison | Morality and Monopoly: The Constitutional
Political Economy of Religious Rules |
| Daniel B. Klein and
Jason Richner | In Defense of the Credit Bureau |
| David E. Dowall | Benefits of Minimal Land-Use Regulations
in Developing Countries |
| Jean-Luc Migué | Trade Barriers in the Theory of Instrument Choice |
| Clark Wiseman | Government and Recycling: Are We Promoting Waste? |
| Svetozar Pejovich | Why Has the Labor-Managed Firm Failed? |

COMMUNICATIONS

- | | |
|--------------------|--|
| W. William Woolsey | The Search for Macroeconomic Stability:
Comment on Sumner |
| Scott Sumner | Index Future Convertibility: Reply to Woolsey |
| Peter A. Schulkin | A Tough Pump to Prime |

VOLUME 12 NUMBER 2 • PUBLISHED BY THE CATO INSTITUTE

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

the CATO JOURNAL

VOLUME 12 NUMBER 2 • PUBLISHED BY THE CATO INSTITUTE • FALL 1992

Editor
JAMES A. DORN

Managing Editor
MONIKA VANBERG

Book Review Editor
SHELDON RICHMAN

EDITORIAL BOARD

TERRY L. ANDERSON
Montana State University

PETER H. ARANSON
Emory University

CHARLES W. BAIRD
California State University
at Hayward

T. E. BORCHERDING
Claremont Graduate School

YALE BROZEN
University of Chicago

JAMES M. BUCHANAN
Center for Study of
Public Choice
George Mason University

RICHARD A. EPSTEIN
University of Chicago

MARILYN R. FLOWERS
Ball State University

JAMES GWARTNEY
Florida State University

STEVE H. HANKE
Johns Hopkins University

ROBERT HIGGS
Seattle University

JERRY L. JORDAN
Federal Reserve Bank of
Cleveland

LEONARD P. LIGGIO
Institute for
Humane Studies

ANTONIO MARTINO
University of Rome

DONALD N. McCLOSKEY
University of Iowa

RICHARD B. MCKENZIE
University of California, Irvine

WILLIAM A. NISKANEN
Cato Institute

SAM PELTZMAN
University of Chicago

GEORGE A. SELGIN
University of Georgia

JUDY SHELTON
Hoover Institution

VERNON L. SMITH
University of Arizona

ROLAND VAUBEL
Mannheim University

LAWRENCE H. WHITE
University of Georgia

LELAND B. YEAGER
Auburn University

The CATO JOURNAL (ISSN 0273-3072) is published in the spring/summer, fall, and winter by the Cato Institute, 1000 Massachusetts Ave., NW, Washington, D.C. 20001-5403. Annual subscriptions are \$24.00 for individuals and \$50.00 for institutions. Single copies are \$8.00. Two-year and three-year subscriptions are \$48.00 and \$72.00, respectively, for individuals, and \$100.00 and \$150.00, respectively, for institutions. Foreign subscribers should add \$5.00 per year for regular delivery and \$10.00 per year for airmail delivery.

Correspondence regarding subscriptions, changes of address, procuring back issues, advertising and marketing matters, etc., should be addressed to the Publications Department. All other correspondence, including requests to quote or reproduce material, should be addressed to the editor.

Unsolicited manuscripts cannot be returned and will be acknowledged only if accompanied by a stamped, self-addressed envelope. There is no submission fee. Articles should focus on specific public policy issues. Accepted manuscripts must conform to the CATO JOURNAL's style requirements. A style sheet is available on request.

Authors should submit three copies of their papers or communications to the editor. An abstract of no more than 250 words should also be included to aid referees. Manuscripts should have wide margins and be double-spaced throughout, including footnotes.

Unsolicited book reviews cannot be returned and will be acknowledged only if they are accepted for publication. They should conform to the above requirements and should provide an in-depth analysis of major studies of public policy. Reviews should range from 800 to 1,000 words.

The views expressed by the authors of the articles are their own and are not attributable to the editor, the editorial board, or the Cato Institute.

The CATO JOURNAL is indexed and abstracted in the *Journal of Economic Literature*, *Current Contents/Social and Behavioral Sciences*, *Social Sciences Citation Index*, *America: History and Life*, *Historical Abstracts*, and the ABI/INFORM data base. The CATO JOURNAL is available on microfilm from University Microfilms International, Ann Arbor, MI 48106.

Printed in the United States of America.

Copyright © 1993 by the Cato Institute.
All rights reserved.

ARTICLES

Anton D. Lowenberg

A Post-Apartheid Constitution for South Africa: Lessons
from Public Choice 297

Peter J. Hill

Environmental Problems under Socialism 321

Marilyn R. Flowers

The Political Economy of Mandated Spending 337

Kenneth W. Costello and David Haarmeyer

Reforming the Bonneville Power Administration 349

Gary M. Anderson and Robert D. Tollison

Morality and Monopoly: The Constitutional Political
Economy of Religious Rules 373

Daniel B. Klein and Jason Richner

In Defense of the Credit Bureau 393

David E. Dowall

Benefits of Minimal Land-Use Regulations
in Developing Countries 413

Jean-Luc Migué

Trade Barriers in the Theory of Instrument Choice 425

Clark Wiseman

Government and Recycling: Are We Promoting Waste? 443

Svetozar Pejovich

Why Has the Labor-Managed Firm Failed? 461

COMMUNICATIONS

W. William Woolsey The Search for Macroeconomic Stability: Comment on Sumner	475
Scott Sumner Index Future Convertibility: Reply to Woolsey	487
Peter A. Schulkin A Tough Pump to Prime	493

BOOK REVIEWS

Generational Accounting: Knowing Who Pays, and When, for What We Spend Laurence J. Kotlikoff <i>Reviewed by</i> William A. Niskanen	501
The Development Frontier: Essays in Applied Economics Peter Bauer <i>Reviewed by</i> Simon Rottenberg	503
Constitutional Environments and Economic Growth Gerald W. Scully <i>Reviewed by</i> Paul H. Rubin	505

A POST-APARTHEID CONSTITUTION FOR SOUTH AFRICA: LESSONS FROM PUBLIC CHOICE

Anton D. Lowenberg

Introduction

As South Africa continues to strive toward constitutional reform, it seems appropriate to examine normative public choice theory for insights that might be relevant to the South African constitutional process. Normative public choice, often called constitutional economics, comprises a comparative analysis of the rules and institutions that govern economic, social and political interactions. The purpose of constitutional economics, in the tradition of Knut Wicksell and James Buchanan, is to identify the attributes of socially efficient rules.¹

Constitutional economics extends the market exchange paradigm to the study of rules and institutions. There is a fundamental distinction between, on the one hand, choice of rules, and, on the other hand, political and economic decisions that are made within a given structure of rules. A central tenet of constitutional economics is that individuals can often agree on a mutually advantageous set of rules, although they normally disagree about specific political and economic questions that are decided within the framework of rules. A level of constitutional deliberation exists conceptually prior to the level of "in-period" politics. While the latter is inherently conflictual, the former bears certain similarities to market exchange if the choice of a constitution reflects voluntary agreement of all parties.

The setting within which constitutional deliberations take place, the "constitutional environment," is an important factor in determining

Cato Journal, Vol.12, No. 2 (Fall 1992). Copyright © Cato Institute. All rights reserved.

The author is Professor of Economics at California State University, Northridge.

¹See Wicksell ([1896] 1958) and Buchanan (1975, 1990).

the efficiency of the constitutional contract that emerges from the deliberations. Recent advances in constitutional economic theory suggest that an efficient contract requires an environment in which all parties are able to exercise an "exit" strategy.² Exit, in this context, refers to the ability of a contracting party to choose the highest valued contractual arrangement from among an array of viable alternatives. Availability of exit opportunities makes the constitutional environment competitive, with the result that the contract produced from this environment is socially efficient.

An efficient contract is mutually beneficial to all participants. It allows individuals to reap the full gains from social cooperation and to maximize the value of post-contractual decisions (Lowenberg and Yu 1990, Gifford 1991). But even an efficient contract will not necessarily survive in the absence of an enforcement mechanism. Individuals have incentives to break rules after the rules have come into effect. This is true even of those individuals who agreed to the rules in the first place. Post-contractual opportunism is especially problematic in the case of political constitutions, which lack effective third-party enforcement.

The problem of enforcement has been addressed in constitutional economics under the rubric of "constitutional maintenance."³ According to the theory of constitutional maintenance, a social contract can be made self-enforcing by including certain types of principles and procedural rules in the constitution. In broad terms, these principles and procedures must entail a commitment to uphold competition in post-constitutional economic and political exchange. The returns to post-contractual opportunism are reduced considerably in a competitive society. Again, the notion of exit provides a useful analogy. In a competitive market, the mechanism through which preferences for private goods are revealed is essentially that of exit (ensured by the availability of alternative trading opportunities of equal or similar value). This market competition paradigm can be extended to the post-constitutional political sphere as well.

The purpose of this paper is to analyze constitutional change in South Africa from the perspective of constitutional economic theory. I will examine the existing constitutional environment in South Africa to determine whether that environment is likely to produce an efficient constitution. I will argue that the existing environment

²See, for example, Lowenberg and Yu (1990, 1992) and the references cited therein.

³Although the term constitutional maintenance is not used by all of these authors, a sampling includes Anderson and Hill (1986), Buchanan and Faith (1987), Wagner and Gwartney (1988), Aranson (1988), Wildavsky (1990), Niskanen (1990), and Lowenberg and Yu (1992).

is not an efficient one, and I will suggest a way to restructure the environment to make it efficient. I also will use constitutional economic theory to identify appropriate constitutional principles that ought to be embodied in a new South African constitution to ensure its survival as a self-enforcing contract. The main conclusion derived from this analysis is that a self-enforcing constitution for South Africa must create decentralization of political authority and provide guarantees of private property rights and freedom of contract.

The Constitution as a Contract⁴

The fundamental problem of social organization can be characterized as a prisoner's dilemma game, in which the dominant strategy, in the absence of agreement or coordination among the players, is noncooperation (for example, stealing, renegeing on the rules, and free riding in contributions to public goods) (Wagner and Gwartney 1988). Although each individual would gain if everyone behaved cooperatively, uncertainty about whether others will cooperate leads to a socially inefficient Cournot-Nash equilibrium of mutual noncooperation.

The only way out of the prisoner's dilemma is an enforceable agreement among the players to cooperate. The agreement must be enforceable because each individual has an incentive subsequently to defect from the cooperative agreement. The mutual gains from social cooperation are a public good, and, as with any other public good, the dominant strategy is to shirk in contributing to its provision. The constitution is essentially a contract intended to secure mutual cooperation, but, without an enforcement mechanism, it cannot succeed in doing so.⁵ Despite the substantial potential gains that a consensual agreement is capable of producing, "the maintenance of such a constitutional contract does not resolve itself naturally" (Wagner and Gwartney 1988, p. 54). Unlike many legal contracts, there is no third-party enforcer, external to the contract, who can ensure that defectors from the cooperative agreement are caught and forced to comply with the terms of the agreement. Of course, many countries have a nominally independent Supreme Court whose purpose is to enforce the constitution. But the Supreme Court can only do this imperfectly in most cases, because the judges themselves are not totally immune from political pressures by groups wishing to subvert the original intent of the constitution (Wagner and Gwartney

⁴The next two sections are based on Lowenberg and Yu (1992).

⁵On the definition and purpose of a political constitution, see Mueller (1991, pp. 326–27).

1988, pp. 40–42; Niskanen 1990, pp. 54–57; Merville and Osborne 1990, p. 40).⁶

Given the unreliability of third-party enforcement, and given the strong individual incentive to defect from social cooperation, a constitutional contract needs to be self-enforcing if it is to be maintained. The first condition necessary for a constitution to be self-enforcing is that it be efficient, i.e., it must provide mutual gains from cooperation. An efficient contract is one that maximizes societal wealth. It produces the largest gains relative to any other feasible contractual arrangement, including the absence of a contract.

The second condition that is necessary for a contract to be self-enforcing is that, in addition to providing mutual gains, recognition of these gains must be sufficient to induce compliance. For example, rules that lower the transaction costs of exchange are mutually beneficial for all participants, and these rules might be expected to be self-enforcing. Most people probably would choose to drive on the same side of the road, even in the absence of traffic police. Such rules are self-enforcing because they solve coordination problems, with the result that all parties have a stake in obeying them (North 1990, p. 41). Many social institutions that are self-enforcing emerge more or less spontaneously, and need very little explicit third-party policing. As an example, Niskanen (1990, p. 58) points out that the mutual desire for continued relations between parties is often enough to enforce contractual performance. A reputation is a valuable asset whenever repeat dealings are involved (North 1990, p. 50). Thus, Axelrod (1984) has demonstrated that cooperation will tend to evolve in repeated prisoner's dilemma games of uncertain duration.

One important attribute of a self-enforcing political constitution is that it must successfully constrain the power of the state. It is not immediately obvious why this should be so, therefore I will elaborate on this point at some length. First, it must be recognized that the rules of a constitution normally are embodied in a state that is given a monopoly of legitimate force. The state is set up to collect taxes to pay for public goods (a necessarily coercive act, given the solution of the prisoner's dilemma game) and to regulate private actions to internalize the spillover effects arising out of externalities and other interdependencies. The state also might be assigned, through the constitution, a role of income redistribution. Individuals at the constitutional stage of deliberation might agree on a mutually beneficial social insurance contract to provide a "safety net" against reductions

⁶Landes and Posner (1975) argue that the U.S. Supreme Court serves the weighted interests of current and past Congresses. See also Aranson (1988, pp. 308–9).

in income due to recession or random events, and therefore might build into the constitution a redistributive function for the state.⁷

Once the instruments of state power are established, however, there is an incentive for groups of individuals to seek to capture these instruments and to use them as vehicles to enrich themselves in ways that are not possible as private citizens. This tendency exists even in liberal democracies (Wagner and Gwartney 1988). In these societies, laws and policies normally are enacted through majority rule. Tullock (1959) has shown that under any voting system that requires less than unanimous approval to implement policies, majority coalitions of interest groups will seek to obtain public provision of special-interest projects. Each of these projects provides large benefits to one group, but is of little interest to anyone else, since the costs are spread over the entire population. Through logrolling, a coalition of interest groups, comprising a bare majority of voters, can get all of their favored projects approved for public provision. Under certain plausible conditions, the total costs of these projects can exceed their total benefits, while cost-spreading through the fisc induces a rational ignorance of this process on the part of the disadvantaged majority (Ordeshook 1986, p. 214; Aranson 1988, p. 291). The result is a diminution of societal welfare through excessive government provision of special-interest legislation for purposes of redistribution (Mueller 1989, p. 84).

A further inefficiency arises as a consequence of rent dissipation occurring through the expenditure of resources by coalitions of special interests seeking redistributive policies. Interest groups spend resources in lobbying politicians and bureaucrats to enact regulations or policies that are beneficial to the groups in question. Politicians respond to these pressures by supplying regulation in accordance with the relative influences of competing interest groups (Peltzman 1976, Becker 1983). Small groups typically are more effective in exerting political influence than large groups, because the benefits obtained by a small group are concentrated among relatively few individuals. This concentration of benefits spurs individual group members to contribute resources and effort to help achieve the group's goals. Small groups are therefore better able to control free riding among their members, which makes them more politically effective than larger groups (Olson 1965). Because the social costs

⁷Ex post redistribution for social insurance purposes, which is potentially wealth enhancing, must be distinguished from inefficient governmental redistribution policies that are primarily a consequence of majority-rule voting and rent seeking. The latter policies are discussed in more detail below. See Mueller (1991, p. 327) for further discussion of governmental redistribution viewed as an insurance contract.

of the policies sought by these politically effective groups are spread over the entire populace, political resistance against these groups' activities is normally weak. Thus, for example, relatively small producer groups often are successful in seeking regulations that raise the costs of entry into their industries, thereby producing rents for existing producers at the expense of large numbers of consumers.

The dominant strategy for any organized interest group in a majoritarian polity is to lobby for policies that create rents for its members and disperse the costs over everyone else (Ordeshook 1986, pp. 213–15). Thus, it is evident that even democratic governments may become instruments of plunder, with those who gain control of the power of the state using that power to prey upon others (Gwartney and Wagner 1988). This brings us back to the Cournot-Nash solution to the prisoner's dilemma game. Although society as a whole would be better off without the existence of redistributational coalitions and the rent seeking that goes along with them, it is privately optimal for all individuals and groups to defect from the cooperative strategy and to join in the plundering. The appropriate notion of government, according to this public choice view, is that of government-as-Leviathan (Wiseman 1990, p. 112).

The contractarian approach to political economy allows the problem of Leviathan to be expressed usefully in terms of principal-agent theory (Anderson and Hill 1986, Merville and Osborne 1990). The political process can be viewed as one in which citizens (principals) grant the power of coercion to an agent (government) in return for the provision of public goods. However, these legitimate public-good outputs of government are like common pool resources, in the sense that no one owns a claim to the capitalized value of public assets whose benefits are widely diffused (Pejovich 1990, p. 33). Political agents soon realize that they can gain greater support by underproviding these public goods and providing instead privatized transfers with concentrated benefits targeted to specific interest groups, financed out of general tax revenues (Anderson and Hill 1986). This leads inevitably to the creation of a Leviathan state. Merville and Osborne (1990) use agency theory to demonstrate formally that, in majority-rule political systems, coalitions of minority factions will induce politicians to systematically break the constitutional contract in order to supply special-interest projects. Political contracts are much more susceptible to this post-contractual opportunism by agents than are private market contracts. In the market context, contractual performance is assured by various self-enforcement mechanisms such as residual claimancy, price signals and competition (Klein and Leffler 1981). All of these serve to lower the

principals' costs of monitoring the agents. The problem of constitutional maintenance in the political arena is essentially a problem of how to replicate some of these mechanisms in order to enforce and monitor the post-contractual performance of political agents.

An efficient constitution, by definition, maximizes the potential gains from cooperative endeavor in the post-constitutional society by constraining noncooperative behavior. However, unless there is some built-in enforcement mechanism, noncooperation will become prevalent and the efficient constitution will be replaced by a Leviathan government, as legislation is increasingly used for redistributive purposes (Wiseman 1990, p. 112; Aranson 1988, p. 310). To prevent this, the constitution must limit the regulatory powers of the state and constrain the scope of state intervention in the realm of private exchange. If the constitution is successful in doing this, the result will be less rent seeking, because a noninterventionist state has fewer rents to award. Rent seeking processes are attenuated in polities where the state is constitutionally limited in its ability to supply special-interest legislation.

The question, then, is how to ensure that a constitution is not only efficient but also self-enforcing. It is not enough for a contract to produce mutual gains from cooperation. That cooperation itself must be maintained. Yet individuals and interest groups inevitably will attempt to subvert the constitutional authority of the state to their own advantage, at the expense of the welfare of the community as a whole. This defection from mutual cooperation diminishes the value of the constitutional agreement and eventually renders meaningless the terms of the constitution. Acceptance of, and adherence to, any contract depends not only on mutuality of gains but also on protection against exploitation of advantage by special-interest groups (Lingle 1991, p. 31). To guarantee this protection, the constitutional contract must be self-enforcing. In the next section, I will consider the conditions necessary for a constitution to be efficient and self-enforcing.

Conditions for Constitutional Efficiency and Maintenance

A constitutional proposal that embodies positively valued prospects for all participants could conceivably garner unanimous consent at the constitutional stage of deliberation (Buchanan 1990, p. 9). It follows that the main normative focus of constitutional economics is on the actual *process* through which rules come to be adopted (Buchanan 1975, p. 6). If that process entails unanimous

consent by all affected parties, then the outcome is efficient, but it is inefficient if elements of coercion are present in the constitutional deliberations.

Thus, unanimous agreement at the stage of constitutional deliberation is a necessary condition for an efficient contract. However, unanimity implies that all parties give their assent knowingly and voluntarily. This, in turn, requires that no one be coerced into the agreement and, most importantly, that all parties possess real, viable alternatives to the contract toward which they are negotiating. A precondition for contractual efficiency, therefore, is the existence of exit options for all contracting parties. An exit option comprises the availability of alternative contracting partners, or alternative and equally-valued contractual arrangements, for individual parties to the constitutional agreement.

The notion of exit is derived from Albert Hirschman's (1970) classic distinction between exit and voice. Exit (and entry) is an important means by which individuals are able to express their preferences, and is precisely the method through which preferences are revealed in competitive markets for private goods. Voice is an alternative method of preference revelation. In this case, the individual expresses his attitudes through complaint, commendation, voting, political influence or some other persuasive device that is not dependent on the threat of exit (Mueller 1989, p. 149).

The essence of the exit strategy in constitutional negotiations is the element of choice among alternative contracting partners. Each individual party to the contract is able to search across an array of potential partners in order to choose the optimal contractual arrangement. No one is coerced to make a contract with anyone who they would not voluntarily choose. The assets brought by any party into a potential cooperative agreement are not specific to any one set of partners or to any one constitutional arrangement. Each party can walk away from an opportunistic partner and seek a higher valued contract with someone else. If exit options were unavailable, however, or if exit were costly, it is unlikely that the resulting contract would be wealth maximizing because individuals involved in contractual negotiations would be locked into dealing with partners whom they would not have chosen if they had the ability to shop around. The availability of exit options helps to defuse potential hold-up threats, bilateral monopoly, and coercion in contractual dealings.⁸

⁸See Alchian, Crawford, and Klein (1978) on the problems of rent appropriation caused by asset specificity and bilateral monopoly. The fewer alternative employment opportunities an asset possesses, the more susceptible it is to post-contractual opportunism. The same principle applies to partners in a political constitution, if they are not brought together initially through a consensual process.

Only with alternatives can agents meet as equals in a trading relationship (Buchanan 1979, p. 34). It is not even necessary that the exit options be exercised, since merely the threat of their use should be enough to restrain rent appropriation. The scope for opportunism is effectively constrained by competition, actual or potential (see Anderson and Hill 1986, pp. 321–22).

Thus, the foregoing discussion implies that the constitutional environment must be characterized by exit options for all parties, if the resulting constitutional contract is to be efficient.⁹ But although an efficient constitution might have considerable potential to create mutual gains for all parties, this is of little value if the rules contained in the constitution are not self-enforcing. As explained in the previous section, constitutional maintenance requires rules that effectively limit the power of the state. As North and Weingast (1989, p. 808) observe, “because constitutional restrictions must be self-enforcing, they must serve to establish a credible commitment by the state to abide by them.” (See also North 1990, p. 50.)

Again, the availability of exit options is an important condition for a self-enforcing constitution, except now the exit options must prevail not only in the constitutional environment but also in the post-constitutional society. Even once all parties have committed to participate by making specific investments in a contractual relationship, it is still possible for exit options to exist post-contractually if the constitution is structured in a way that commits the participants to maintain a competitive environment in the post-constitutional society. In a competitive social setting, each party possesses a credible threat that can be used to prevent other parties from behaving opportunistically. This threat is precisely the option of “exiting” to make a deal with an alternative trading partner.¹⁰ Competition ensures that the terms of trade will not differ significantly from one trading partner to

⁹A familiar proposition in constitutional economics is that an efficient constitution requires a “veil of uncertainty,” in which participants in constitutional deliberations are ignorant of their future interests in post-constitutional distributive struggles (Buchanan and Tullock 1962; Buchanan 1986, p. 95). Unfortunately, it is by no means clear how such uncertainty can be achieved in a real-world setting. However, according to Lowenberg and Yu (1992), the existence of exit opportunities for contracting parties can produce efficient results identical to those that would be produced by a veil of uncertainty. With viable exit options for all parties, no interest group, no matter how effective it might be in exerting political influence, can succeed in extracting terms from another group that are any less favorable to it than the best terms available in alternative cooperative agreements.

¹⁰A self-enforcing constitution embodies a credible commitment to support exchange in the post-constitutional society. This commitment grows out of the ability to credibly threaten exit. On the distinction between credible threats and credible commitments, see Williamson (1989, pp. 16–17).

another.¹¹ Therefore if one partner attempts to appropriate rents from another, the latter can simply refuse to do business with the opportunistic partner and switch to someone else at little or no cost to himself. The threat, "I won't deal with you any more if you don't play by the rules" becomes perfectly credible in this scenario. It follows that a strategy of rent seeking and noncooperation does not pay if the post-constitutional society is sufficiently competitive.

The availability of exit options means that each individual is guaranteed a wide range of choice in many spheres of political and economic life. "When people have other decisionmakers or forums to turn to (alternative firms, alternative courts, alternative legislatures), such options enable private individuals to impose costs (consequences) on those attempting to [behave opportunistically]. . ." (Crew and Twight 1990, p. 27). Constitutional rules guaranteeing these options are inherently self-enforcing because of the competitive characteristics of the society that they create.

A number of constitutional procedures and principles that could help to maintain a competitive, pluralistic environment for political and economic exchange have been proposed in the constitutional economics literature.¹² A common theme running through many of these proposals is the strengthening of regional and local government relative to national government. The existence of separate jurisdictions with some protected powers within a constitutional federation inhibits coercive behavior by the government (Wiseman 1990, pp. 121–22). Such an arrangement facilitates migration at low cost between federal sub-regions and thereby enhances competition between these sub-regions. The ability to "vote with their feet" enables individuals to reveal their preferences by exercising a "spatial" exit option. The resulting mobility forces competitive governmental units to supply public goods in preferred quantities and to "price" them broadly in line with relative marginal evaluations (Buchanan and Faith 1987, p. 1023). This is known as the Tiebout effect (Tiebout 1956), whereby individuals sort themselves across communities in accordance with their preferences for the packages of taxes and public goods provided in each community. To the extent that the Tiebout effect is fully operative, the wealth created through government protection of property rights is not specialized to any one jurisdiction. The ability of the owners of these rights to move

¹¹Mueller (1991, p. 332) points out that international mobility of capital in modern financial markets limits the scope for monopolistic and monopsonistic exploitation of capital or labor, as long as government policy, driven by special-interest pressures, does not intervene.

¹²See Lowenberg and Yu (1992) for a review of some of these proposals.

to competing jurisdictions protects them from potential rent appropriation by a coercive government (Anderson and Hill 1986, p. 323).¹³

Constitutional rules allowing for mobility and political diversity will help to ensure a competitive political system with many alternative forums for seeking political redress. Wildavsky (1990) argues that a federalist constitutional system embodies this kind of diversity and effectively constrains the monopoly power of government. The defining characteristic of federalism is competition. Federalist institutions provide opportunities for experimentation with alternative approaches to legislative action. In dealing with local legislators or local courts, for example, aggrieved parties can essentially vote with their feet by seeking relief elsewhere if they are not satisfied with the ruling of a particular level of government (Crew and Twight 1990, p. 27).¹⁴ Thus, an important element of a self-enforcing constitution is the effective self-enforcing constraint that is placed on the power of the state as a consequence of inter-jurisdictional political competition.

The ability of majority coalitions to use legislation to redistribute income and wealth from the rest of the community can be constrained by constitutionally limiting the redistributive element of legislation through adoption of appropriate electoral and voting rules. A strong government can be produced out of a winner-takes-all system of single-member constituencies, in which a runoff election is held to determine which party will hold an absolute majority in the legislature (Mueller 1991, pp. 334–36). But a disadvantage of this system is the potential alienation of voters in the minority who feel inadequately represented. To avoid this, a proportional representation system might be used to ensure that the composition of the legislature reflects the full range of voter preferences.¹⁵ Under proportional representation, however, it will be necessary to use a supramajority voting rule (two-thirds or three-quarters) in the legislature to mitigate the tendency for logrolling coalitions to subvert the legislative

¹³See Epplé and Romer (1991) for further analysis of the constraints that potential migration, as well as other factors such as property ownership, place on the ability of local governments to undertake income redistribution.

¹⁴Wildavsky (1990, p. 45) provides further elaboration of this process:

Anyone who has to deal with governmental agencies, or who wants to apply for grants, knows that it is useful to have a multiplicity of points of access and largess. It is then possible to play them off against each other or simply to go to the most receptive place. Individuals can thus sort themselves out among agencies with different rules, objectives, and personal predilections. Similarly, grievances encountered in one place may be redressed in another.

¹⁵Proportional representation also can result in higher voter turnout in elections. See Kaempfer and Lowenberg (1993).

process for redistributive purposes. Even under the two-party, winner-takes-all system, which depends on simple majority rule, a supra-majority rule would be desirable for legislation that is overtly redistributive (such as import tariffs), to prevent potential zero-sum or negative-sum redistributions.¹⁶

We can conclude that an efficient constitution requires as a precondition a constitutional environment characterized by availability of exit options in the sense of voluntary participation. But contractual efficiency is not sufficient to ensure contractual maintenance. For an efficient constitution to survive, it must be self-enforcing. An effective way to make a constitution self-enforcing is to build in procedures and institutions designed to establish a competitive post-constitutional environment for political and economic exchange. Constitutional principles that fulfill this purpose include guaranteed property rights, freedom of association, freedom of migration, local political autonomy, rights of communities or regions to secede from a polity or incorporate as separate entities within a polity, use of supramajority voting rules in legislative bodies, etc.¹⁷ Some of these will be examined in more concrete terms in the next section, which uses public choice analysis to generate proposals for restructuring the South African constitutional environment, and makes suggestions for specific principles that ought to be included in a new South African constitution.

Reinventing Constitutional Discourse in South Africa

Constitutional economics suggests that there are two difficulties associated with efficient constitution formation and maintenance in the South African context. First, the present constitutional environment does not offer realistic exit opportunities for all participants. Second, the political heritage of the major players in the South African constitutional process does not include the traditions of competitive exchange that are needed for constitutional maintenance. Consideration of these two general areas of difficulty leads to the following proposals for alteration of the current constitutional process:

The Constitutional Environment

Currently, each of the major participants in the constitutional process is seeking exclusive control, or at least a substantial share in

¹⁶This argument is made by Mueller (1991, p. 343).

¹⁷For more detailed discussion of constitutional principles that are needed to establish a society based on private rights, see Mueller (1991, pp. 328–33).

control, of a unitary polity. Both the African National Congress (ANC) and the National Party represent broad alliances of whites and non-whites. These two alliances are competing for leadership in the government of a centralized South African state. All ethnic, language and ideological groups, within the two major coalitions and outside them, are locked into the necessity of making a constitutional arrangement with one another to live together in a single society. There are no exit options.¹⁸

Exit opportunities could be introduced into the South African constitutional environment by restructuring constitutional discourse. The restructuring would entail creation, *ab initio*, of separate constitutional destinies for each group within the South African polity. These separate constitutional futures could be compromised or abandoned only with the consent of the groups in question. Thus each group would enter into constitutional negotiations *de novo*, with realistic options of choosing alternative partners, or no partners at all, with whom to join in a cooperative social arrangement.

This constitutional environment is a realizable prospect because many of the major political groups involved in the South African constitutional process are, or could be, defined in terms of their territorial identities. Most of the ANC's supporters, for example, are Xhosas from the Transkei and the urban townships of the Witwatersrand. Most of the supporters of the Inkatha Freedom Party are Zulus from Natal. The extreme right-wing parties draw most of their support from whites in the northern Transvaal. Coloreds (mixed-race) people are concentrated in the western Cape province, and Indians are concentrated in Natal. Urban areas such as Soweto, that are ethnically and politically mixed, could enter into the constitutional process as separate city-states.

The notion of a South Africa comprising a collection of autonomous groups is not alien to the constitutional debate. South Africa already has considerable experience with political decentralization and restructuring due to the "homelands" policy of the National Party government (Kendall and Louw 1987, p. 192). Almost all the major parties to the South African conflict now endorse, in one form or another, a federal, consociational or partition constitution, or, at a minimum, some degree of constitutional protection for group identi-

¹⁸The February 1993 agreement between the ANC and the National Party makes clear that, initially, political power is to be centralized in a single national government, and any consideration of regional autonomy will be deferred until after the election of a national constituent assembly (as reported by Bill Keller in the *New York Times*, 19 February 1993, pp. A1, A4).

ties.¹⁹ According to Kendall and Louw (1987, pp. 98–99, 110), the National Party, the ANC, and the Democratic Party support some form of devolution or recognition for national groups, although the ANC's Freedom Charter still calls for a unitary state (see Lingle 1991, p. 32). The right-wing Conservative Party and Afrikaner Weers-tandsbeweging (AWB) support partition.

The parties involved in constitutional deliberations need to consider a number of factors that together determine the optimal number and configuration of polities. These factors include, among other things, the geographic distribution of the population, physical characteristics of the country, costs of running separate governmental units, and the benefits of matching the suppliers of public goods to the group benefiting from them the most.²⁰

The advantage of placing constitutional discourse in a decentralized federal framework is that each group would be assumed from the outset to have a right to an autonomous polity, defined territorially, and no group would be in a position of competing for political power over a unitary state. The ability to shop around among alternative partners and alternative arrangements among autonomous groups provides the basis for efficient constitutional outcomes. Moreover, constitutional exit options reduce conflicts at the level of the national constitution by allowing regional and other lower levels of government to define particular individual rights in their own constitutions, especially in those situations where agreement at the national level is not possible (Mueller 1991, p. 342).

Federalism

As pointed out above, in the absence of a self-enforcing constraint on the scope of government regulation, a democratic majoritarian polity is likely to produce a Leviathan state that uses its regulatory powers to make inefficient redistributions. Government regulation emasculates competitive exchange, in both economic and political spheres, and replaces it with monopoly rents.

To constrain Leviathan, decentralization of political power is desirable. Federalism within a unified (as opposed to unitary)²¹ state

¹⁹Consociationalism refers to the decentralized "canton" system of government epitomized by Switzerland, but also present at various times in other ethnically or linguistically diverse countries such as Belgium, the Netherlands and Austria (Lijphart 1985). On consociational proposals for South Africa, see Schlemmer (1978), Wiechers (1978), Hanf, Weiland, and Vierdag (1981), Boule (1984), Lijphart (1985, pp. 31–35), Kendall and Louw (1987), and Giliomee and Schlemmer (1989).

²⁰See Mueller (1991, p. 340), Tullock (1969), and Oates (1972).

²¹On this distinction, see Lingle (1991, p. 39).

provides a viable approach to political decentralization. Wildavsky (1990, p. 41) argues that the appropriate term is "noncentralization." This refers to independent centers of power in geographic areas that differentiate and compete among themselves. Noncentralization, according to Wildavsky, is characteristic of federalism, whereas decentralization could be consistent with delegation of authority within a unitary polity.

A federalist constitution possesses the advantage of entrenching mobility and spatial exit opportunities that are needed to facilitate Tiebout competition and efficient public good provision. As pointed out above, the most important attribute of federalism is competition. "The operational meaning of federalism is found in the degree to which the constituent units disagree about what should be done, who should do it, and how it should be carried out" (Wildavsky 1990, p. 43). Wildavsky adds that a true federalist constitution should extend competition between jurisdictions from the state level to regional and local levels of government as well (1990, p. 48).

An additional advantage of a federalist structure arises in situations where ethnic, religious and language groups reside in particular sections of a country. Under federalism, "conflicts over distributional and rights issues can be reduced, thus freeing each polity for the task of discovering and supplying a community's collective goods by drawing federalist political boundaries to separate the cultural groups that feel hostility toward one another" (Mueller 1991, pp. 340-41). The responsibility of each level of government should be defined by the magnitude of the spillovers from production and consumption of local public goods. For example, the geographic area impacted by the provision of a fire station or trash collection service is relatively small and could be administered by a town or city. However, externalities that affect larger areas, such as traffic on a highway or a river, require regional levels of government (Mueller 1991, pp. 339-40).

An important insight derived from recent work in comparative institutional analysis is that even informal constitutional arrangements have considerable impacts on economic performance (North 1989, 1990). Many scholars have argued that a decentralized constitution is a precondition for successful economic development. For example, according to Rosenberg and Birdzell (1986), the dynamics of competition among a multiplicity of independent decision centers helped to fuel Western growth. North (1990, p. 130) likewise argues that competition among political units was an important reason for successful European economic performance in contrast to China, Islam, and other more centralized systems. Vanberg (1992) points out

that the competitive-evolutionary process that facilitates economic growth and development requires an environment characterized by diffusion of authority and decentralization. Moreover, as Brennan and Buchanan (1980, pp. 184–85) demonstrate, decentralized government can substitute for explicit constitutional constraints on the power to tax and spend (see also Oates 1989).

In the South African context, the maintenance of post-constitutional exit options in the form of local political autonomy is just as important as the availability of exit options at the constitutional stage of deliberation. A new South African constitution should contain rules entrenching the autonomy of local governments. One way to ensure the diffusion of power from central government to regional and local governments is to have procedural rules requiring larger legislative majorities for legislation at higher levels of government (Wagner and Gwartney 1988).²² Such rules would raise the cost of using central government legislation for redistributive purposes. The coercive power of any one governmental authority or jurisdiction in post-apartheid South Africa can be further constrained by permitting free migration of individuals and a generalized right of residence. As pointed out above, inclusion of these principles in the constitution creates Tiebout competition which diminishes the scope for rent appropriation brought about through redistributive legislation.

The South African constitution also should guarantee the rights of disaffected communities to secede from an existing political jurisdiction,²³ and the right of incorporation of new jurisdictions, such as cities, within larger regional entities. Buchanan and Faith (1987) use the term “internal exit” to refer to the establishment of new political units by individuals defecting from a given polity. Internal exit, like other forms of spatial competition, imposes effective constraints on the potentially exploitative behavior of a dominant political coalition.

Economic Rights

In addition to federalism, constitutional maintenance requires a commitment to uphold *individual* rights to property and contract. In a society in which competitive exchange is the norm and monopoly the exception, rent seeking processes are substantially attenuated. Private monopolies, most of whom’s long-term survival depends on state regulation of private markets, can be inhibited by procedural

²²On the advantages of procedural rules over substantive constraints on government, see Wagner and Gwartney (1988), and Dorn (1991, p. 164).

²³In a plural society, it is often desirable for a segmentally homogeneous territory to secede from the federation (Lijphart 1985, p. 99n).

rules protecting property rights and freedom of exchange from incursions of government laws. For example, legislative redistribution can be reduced by requiring a supramajority vote for all legislation that is not explicitly aimed at deciding on efficient levels of public goods or externalities. Mueller (1991, pp. 329–30), for example, advocates that supramajorities be required for any legislation enabling the government to intervene in a market by restricting prices, quantities or entry, or to engage directly in the production and sale of a good, or to tax or regulate foreign trade.

As pointed out above, the theoretical basis for unrestrained market exchange is that exchange provides post-constitutional exit opportunities for all individuals in the form of alternative trading partners. These exit options help to reduce opportunistic or noncooperative strategies in the post-constitutional society.

South Africa unfortunately faces formidable barriers to the establishment of a free enterprise society based on private rights. The dominant political ideologies, both those of the National Party and the Africanist parties,²⁴ were formulated in a statist tradition. For decades, the South African political economy has been characterized by heavy-handed state regulation of private exchange. In order to support a massive affirmative action program for white workers and farmers, the apartheid state intervened in almost every aspect of social and economic life, severely circumscribing private property rights and curtailing individual freedom to contract (Hazlett 1988, Lowenberg 1989). The result was an economic system in which rents were allocated through the political process and wealth was often systematically destroyed by removing individual incentives to allocate resources efficiently. The South African economy remains one of the most socialized in the noncommunist world, with many key industries owned by the state and a large proportion of the labor force employed in giant state bureaucracies.²⁵

Many South Africans, especially among the younger generation of blacks, believe that the apartheid economy was a capitalist one. Consequently, they believe that the post-apartheid society must abandon capitalism. In fact, South Africa never bore much resemblance to a liberal, free-enterprise system. In large part, the capitalism

²⁴These include the ANC and the more radical Pan Africanist Congress. For detailed descriptions of the policies and ideologies of the major participants in the South African constitutional process, see Kotze (1990).

²⁵Until recently, state-owned firms monopolized the production of electricity, iron and steel, railroad transportation and broadcasting among many other goods and services. In 1986 the ratio of public sector employment to total employment was 30 percent. See Kendall and Louw (1987, pp. 60–61, 72).

that exists in South Africa is more akin to the state capitalism of the former Soviet Union than it is to that of Western capitalist countries (Lingle 1990). Large monopolies, both privately and publicly owned, are protected by regulations that raise the costs of entry by smaller enterprises. Throughout the heyday of apartheid, influx controls and residential segregation inhibited the freedom of resource owners to choose efficient combinations of factors of production. As in the former Soviet Union, true free enterprise has flourished only in black markets or informal sectors of the economy.

There is a strong temptation among members of the Africanist parties to design a post-apartheid constitution along socialist lines, creating an omnipotent, centralized and unitary state and entrusting to it the job of confiscating property and redistributing wealth to those who have suffered so unjustly under apartheid. Given the historical heritage of apartheid and the struggle against apartheid, it is not surprising that, as Lingle (1991) points out, both the National Party and the ANC share a collectivist vision of government, in which political power is centralized and directed toward redistribution favoring special-interest groups. The only difference, of course, is in the identities of the groups who would benefit from the ANC's policies as opposed to those of the National Party. In fact, there is a remarkable resemblance between the economic system that prevailed under apartheid and the extensive state ownership and large-scale state intervention envisaged in the Freedom Charter to which the ANC subscribes (Lingle 1991, p. 40). Vorhies (1991, p. 13) argues that, in its continued support for strong centralized government and nationalized monopolies, the ANC is a virtual clone of the National Party.

Constitutional economics provides a strong argument against pursuing this statist model. To preserve the cooperative benefits of the constitution, individuals should be guaranteed exit options in all political and economic interactions. In the economic sphere, this translates into a system of private property rights and minimal restrictions on exchange. Again, procedural rules built into the constitution that constrain the regulatory powers of government, especially central government, can help to maintain such a system.

Conclusion

Constitutional efficiency requires voluntary unanimous agreement. But for an agreement to be voluntary, the parties must possess viable alternative options that they willingly agree to give up in order to commit themselves to one constitution out of a set of possible

constitutions to which they might otherwise commit. Therefore an efficient social contract can only emerge out of a constitutional environment that provides exit options for all contracting parties. Exit options allow each party a choice of partners with whom to make cooperative arrangements of more or less equal value. The availability of exit opportunities removes all elements of coercion from the constitutional bargaining process, and thereby ensures that the constitutional outcome is the highest valued of all arrangements that are potentially achievable.

Even if a mutually beneficial cooperative arrangement does emerge, however, it will not necessarily survive unless the constitution is self-enforcing. There are strong incentives for individuals to defect from cooperative strategies in the post-contractual setting. To limit these tendencies, the post-constitutional society should contain many opportunities for political and economic exchange. Only the fluidity of competition can ensure compliance with the cooperative agreement embodied in the constitution. Competition closes off avenues for opportunism and rent seeking. Therefore the constitution should entrench rules that effectively restrain the regulatory and redistributive powers of the state.

The constitutional process currently under way in South Africa is unlikely to produce an efficient constitution, unless the existing constitutional environment is replaced with one that is able to offer viable exit options to all participants. To achieve this, the format of the constitutional debate should be restructured into a voluntary association of autonomous groups, each seeking to form a cooperative society by choosing from an array of possible constitutional arrangements. This restructuring of constitutional discourse could help to ensure that an efficient constitutional contract (or contracts) eventually will emerge out of the current process. By definition, an efficient constitution is a positive-sum agreement. Therefore, an efficient social contract enables redistribution of wealth to be achieved without destroying wealth. Slices of a growing pie can be reapportioned without necessarily making anyone worse off in absolute terms. Simply by abolishing the inefficiencies of apartheid, an efficient constitution has the potential to make both blacks and whites better off while also achieving a redistribution in favor of blacks (Kendall and Louw 1987, p. 86).

But this is not enough. The new constitution must be made self-enforcing to avoid the pitfalls of the past. The apartheid edifice crumbled because it was not a self-enforcing constitutional system. It was never the product of a consensual process of constitutional deliberation, nor did it provide a competitive environment for social

interaction. Instead it was a system of political plunder orchestrated through a large, interventionist state. While blacks were the main losers, consumers and taxpayers of all colors paid a high price to support monopolized production, emasculated trade, and the appetite of a Leviathan state.

Several elements are necessary to ensure that a new post-apartheid constitution is self-enforcing. Such a constitution ought to provide for decentralization and plurality of political power through the adoption of a federal system. Procedural rules designed to diffuse political power from the central government to regional and local levels of government will help to create spatial competition among jurisdictions and efficient provision of governmental services. The redistributive tendencies of majoritarian government should be offset with the use of supramajority legislative voting rules. In addition, the principles contained in a new South African constitution should establish a free enterprise society based on private property rights and a minimalist state. A post-apartheid constitution can be made self-enforcing only by building in a commitment to political and economic competition.

References

- Alchian, Armen A.; Crawford, Robert G.; and Klein, Benjamin. "Vertical Integration, Appropriable Rents, and the Competitive Contracting Process." *Journal of Law and Economics* 21 (October 1978): 297-326.
- Anderson, Terry L., and Hill, Peter J. "Constraining the Transfer Society: Constitutional and Moral Dimensions." *Cato Journal* 6 (Spring/Summer 1986): 317-39.
- Aranson, Peter H. "Procedural and Substantive Constitutional Protection of Economic Liberties." In *Public Choice and Constitutional Economics*, pp. 285-313. Edited by James D. Gwartney and Richard E. Wagner. Greenwich, Conn.: JAI Press, 1988.
- Axelrod, Robert M. *The Evolution of Cooperation*. New York, N.Y.: Basic Books, 1984.
- Becker, Gary S. "A Theory of Competition among Pressure Groups for Political Influence." *Quarterly Journal of Economics* 98 (August 1983): 371-400.
- Boulle, Laurence J. *Constitutional Reform and the Apartheid State: Legitimacy, Consociationalism and Control in South Africa*. New York, N.Y.: St. Martin's Press, 1984.
- Brennan, Geoffrey, and Buchanan, James M. *The Power to Tax: Analytical Foundations of a Fiscal Constitution*. Cambridge: Cambridge University Press, 1980.
- Buchanan, James M. *The Limits of Liberty: Between Anarchy and Leviathan*. Chicago, Ill.: University of Chicago Press, 1975.
- Buchanan, James M. *What Should Economists Do?* Indianapolis, Ind.: Liberty Press, 1979.

- Buchanan, James M. "Rights, Efficiency, and Exchange: The Irrelevance of Transactions Cost." In James M. Buchanan, *Liberty, Market, and State: Political Economy in the 1980s*, pp. 92–107. New York, N.Y.: New York University Press, 1986.
- Buchanan, James M. "The Domain of Constitutional Economics." *Constitutional Political Economy* 1 (Winter 1990): 1–18.
- Buchanan, James, M., and Faith, Roger L. "Secession and the Limits of Taxation: Toward a Theory of Internal Exit." *American Economic Review* 77 (December 1987): 1023–31.
- Buchanan, James M., and Tullock, Gordon. *The Calculus of Consent*. Ann Arbor, Mich.: University of Michigan Press, 1962.
- Crew, Michael A., and Twight, Charlotte. "On the Efficiency of Law: A Public Choice Perspective." *Public Choice* 66 (July 1990): 15–36.
- Dorn, James. A. "Madison's Constitutional Political Economy: Principles for a Liberal Order." *Constitutional Political Economy* 2 (Spring/Summer 1991): 163–86.
- Epplé, Dennis, and Romer, Thomas. "Mobility and Redistribution." *Journal of Political Economy* 99 (August 1991): 828–58.
- Gifford, Adam Jr. "A Constitutional Interpretation of the Firm." *Public Choice* 68 (January 1991): 91–106.
- Giliomee, Hermann, and Schlemmer, Lawrence. "Review: Guidelines and Goals for Resolution." In *Negotiating South Africa's Future*, pp. 146–55. Edited by Hermann Giliomee and Lawrence Schlemmer. Johannesburg: Southern Book Publishers, 1989.
- Hanf, Theodor; Weiland, Heribert; and Vierdag, Gerda. *South Africa: The Prospects of Peaceful Change*. London: Rex Collings, 1981.
- Hazlett, Thomas W. "Economic Origins of Apartheid." *Contemporary Policy Issues* 6 (October 1988): 85–104.
- Hirschman, Albert O. *Exit, Voice, and Loyalty*. Cambridge, Mass.: Harvard University Press, 1970.
- Kaempfer, William H., and Lowenberg, Anton D. "A Threshold Model of Electoral Policy and Voter Turnout." *Rationality and Society* 5 (January 1993): 107–26.
- Keller, Bill. "Mandela's Group Accepts Five Years of Power-Sharing." *New York Times*, 19 February 1993, pp. A1, A4.
- Kendall, Frances, and Louw, Leon. *After Apartheid: The Solution for South Africa*. San Francisco, Calif.: Institute for Contemporary Studies, 1987.
- Klein, Benjamin, and Leffler, Keith B. "The Role of Market Forces in Assuring Contractual Performance." *Journal of Political Economy* 89 (August 1981): 615–41.
- Kotze, Hennie. "The Stakeholders." In *The Myth Makers: The Elusive Bargain for South Africa's Future*, pp. 36–62. Edited by Pierre du Toit and Willie Esterhuysen. Pretoria: Southern Book Publishers, 1990.
- Landes, William M., and Posner, Richard A. "The Independent Judiciary in an Interest-Group Perspective." *Journal of Law and Economics* 18 (December 1975): 875–901.
- Lijphart, Arend. *Power Sharing in South Africa*. Berkeley, Calif.: Institute of International Studies, 1985.
- Lingle, Christopher. "Apartheid as Racial Socialism." *Kyklos* 43 (2) (1990): 229–47.

- Lingle, Christopher. "Constitutional Political Economy and Analysis of Reform Alternatives for Post-Apartheid South Africa." *Journal for Studies in Economics and Econometrics* 15 (1) (1991): 31-42.
- Lowenberg, Anton D. "An Economic Theory of Apartheid." *Economic Inquiry* 27 (January 1989): 57-74.
- Lowenberg, Anton D., and Yu, Ben T. "Constitutional Environments and the Contractual State: The Cases of South Africa and Hong Kong." *Journal of Comparative Economics* 14 (March 1990): 51-69.
- Lowenberg, Anton D., and Yu, Ben T. "Efficient Constitution Formation and Maintenance: The Role of 'Exit'." *Constitutional Political Economy* 3 (Winter 1992): 51-72.
- Merville, Larry J., and Osborne, Dale K. "Constitutional Democracy and the Theory of Agency." *Constitutional Political Economy* 1 (Fall 1990): 21-47.
- Mueller, Dennis C. *Public Choice II*. Cambridge: Cambridge University Press, 1989.
- Mueller, Dennis C. "Choosing a Constitution in East Europe: Lessons from Public Choice." *Journal of Comparative Economics* 15 (June 1991): 325-48.
- Niskanen, William A. "Conditions Affecting the Survival of Constitutional Rules." *Constitutional Political Economy* 1 (Spring/Summer 1990): 53-62.
- North, Douglass C. "A Transaction Cost Approach to the Historical Development of Politics and Economies." *Journal of Institutional and Theoretical Economics* 145 (December 1989): 661-68.
- North, Douglass C. *Institutions, Institutional Change and Economic Performance*. Cambridge: Cambridge University Press, 1990.
- North, Douglass C., and Weingast, Barry R. "Constitutions and Commitment: The Evolution of Institutions Governing Public Choice in Seventeenth-Century England." *Journal of Economic History* 49 (December 1989): 803-32.
- Oates, Wallace E. *Fiscal Federalism*. London: Harcourt Brace, 1972.
- Oates, Wallace E. "Searching for Leviathan: A Reply and Some Further Reflections." *American Economic Review* 79 (June 1989): 578-83.
- Olson, Mancur. *The Logic of Collective Action: Public Goods and the Theory of Groups*. Cambridge, Mass.: Harvard University Press, 1965.
- Ordeshook, Peter C. *Game Theory and Political Theory*. Cambridge: Cambridge University Press, 1986.
- Pejovich, Svetozar. *The Economics of Property Rights: Towards a Theory of Comparative Systems*. Dordrecht: Kluwer Academic Publishers, 1990.
- Peltzman, Sam. "Toward a More General Theory of Regulation." *Journal of Law and Economics* 19 (August 1976): 211-40.
- Rosenberg, Nathan, and Birdzell, L. E. Jr. *How the West Grew Rich: The Economic Transformation of the Industrial World*. New York, N.Y.: Basic Books, 1986.
- Schlemmer, Lawrence L. "Social Implications of Constitutional Alternatives in South Africa." In *Constitutional Change in South Africa*, pp. 258-75. Edited by John Benyon. Pietermaritzburg: University of Natal Press, 1978.
- Tiebout, Charles M. "A Pure Theory of Local Expenditures." *Journal of Political Economy* 64 (October 1956): 416-24.

- Tullock, Gordon. "Some Problems of Majority Voting." *Journal of Political Economy* 67 (December 1959): 571-79.
- Tullock, Gordon. "Federalism: Problems of Scale." *Public Choice* 6 (Spring 1969): 19-30.
- Vanberg, Viktor. "Innovation, Cultural Evolution and Economic Growth." In *Explaining Process and Change: Approaches to Evolutionary Economics*, pp. 105-121. Edited by Ulrich Witt. Ann Arbor, Mich.: University of Michigan Press, 1992.
- Vorhies, Frank. "The Mercantilist Roots of Apartheid." Working paper. Department of Business Economics, University of the Witwatersrand, Johannesburg, June 1991.
- Wagner, Richard E., and Gwartney, James D. "Public Choice and Constitutional Order." In *Public Choice and Constitutional Economics*, pp. 29-56. Edited by James D. Gwartney and Richard E. Wagner. Greenwich, Conn.: JAI Press, 1988.
- Wicksell, Knut. "A New Principle of Just Taxation." 1896. Reprinted in *Classics in the Theory of Public Finance*, pp. 72-118. Edited by Richard A. Musgrave and Alan T. Peacock. London: Macmillan, 1958.
- Wiechers, Marinus. "Possible Structural Divisions of Power in South Africa." In *Constitutional Change in South Africa*, pp. 107-18. Edited by John Benyon. Pietermaritzburg: University of Natal Press, 1978.
- Wildavsky, Aaron. "A Double Security: Federalism as Competition." *Cato Journal* 10 (Spring/Summer 1990): 39-58.
- Williamson, Oliver E. "Operationalizing the New Institutional Economics: The Transaction Cost Economics Perspective." Working paper no. BPP-38, Working Paper Series in Business and Public Policy. Center for Research in Management, Berkeley Business School, University of California, Berkeley, January 1989.
- Wiseman, Jack. "Principles of Political Economy: An Outline Proposal, Illustrated by Application to Fiscal Federalism." *Constitutional Political Economy* 1 (Winter 1990): 101-24.

ENVIRONMENTAL PROBLEMS UNDER SOCIALISM

Peter J. Hill

The Problem

For several decades, the degree to which socialism generates environmental problems has been a matter of controversy. Oskar Lange, writing in the 1930s, argued that socialism would do a much better job of accounting for all costs than capitalism; hence externalities would be minimized (Lange 1938). On the other hand, some scholars have argued that central planning and state ownership of the means of production do not necessarily prevent environmental degradation (Goldman 1972). However, with the collapse of the Soviet Union and the coming of democracy to Eastern Europe, information has flowed much more freely, and the extent of ecological disruption has become more widely known. As reports have accumulated, it has become clear that the environmental problems of these countries are large indeed.¹

Children from the Upper Silesia area of Poland have been found to have five times more lead in their blood than children from Western European cities. Half of the children in that area suffer from pollution related illnesses (Maremont, Kapstein, and Schares 1990). Romania has no air pollution controls and in some areas horses are only allowed to stay for two or three years. "They have to be taken away, or else they will die," says Dr. Alexandru Balin, a physician in the Romanian town of Copsa Mica (Painton 1990, p. 40).

The worst air pollution is in the industrial corridor of the southern part of East Germany, across northern Czechoslovakia, and into

Cato Journal, Vol. 12, No. 2 (Fall 1992). Copyright © Cato Institute. All rights reserved.

The author is the George F. Bennett Professor of Economics at Wheaton College. He thanks Terry Anderson and Jim Halteman for helpful comments.

¹Although there are many sources that document the environmental crises of the ex-socialist countries, the most complete discussion to date is found in Feshbach and Friendly (1992).

southern Poland (Simons 1990). In this area a soft lignite coal with a high sulfur content is the main fuel source. In Leuna, in what was formerly East Germany, at any given time 60 percent of the population suffers from respiratory ailments (Charles 1990). Four out of five children in Espenhain develop chronic bronchitis or heart ailments by the age of seven (Painton 1990). In Telpice, a town in northwest Czechoslovakia, air pollution keeps children inside for about a third of the winter. In an attempt to preserve the children's health, school is held in cleaner towns six weeks a year (Maremont, Kapstein, and Schares 1990).

Water pollution has also been a significant problem in numerous Eastern European countries. Drinking water in Hungary is seriously contaminated with arsenic. Sewage treatment is nonexistent or very primitive in many large cities. Bulgarian agriculture suffers from heavy metals pollution through irrigation waters of much of its best farming regions (Maremont, Kapstein, and Schares 1990).

As deplorable as conditions are in Eastern Europe, the situation in the former Soviet Union is little better. Air and water pollution abound there also. Even more significant has been the massive alteration of two major bodies of water, Lake Baikal and the Aral Sea.

Lake Baikal, the largest and deepest fresh water lake on earth, had long been noted for its remarkably purity. The Lake Baikal basin is also home to over 1,200 fish and plant species not found elsewhere (Feshbach and Friendly 1992, p. 117). In 1957, Soviet planners decided paper mills should be located on the shores of Lake Baikal to take advantage of the surrounding forests and ample supply of clean water. Massive pollution resulted and significant deterioration of the lake occurred. The effluent from the pulp factory is discharged directly into the lake and has created a polluted zone 23 miles wide (Thompson 1989).

Even more dramatic has been the destruction of the Aral Sea. An enormous, shallow, body of water located in the south-central portion of the Soviet Union, it was once larger than any of the Great Lakes except Lake Superior (Micklin 1988). Shortly after the Communists came to power a decision was made that the Soviet Union should become self-sufficient in cotton production. This required massive diversions of irrigation water, and most of it has come from the two rivers that feed the Aral, the Amu Darya and the Syr Darya. The inflow to the Aral Sea has been reduced to almost nothing and as a consequence the Aral Sea has been decreasing rapidly in size. From 1960 to the present the area of the Aral Sea has diminished 40 percent and the volume 66 percent (Micklin 1988, p. 1170).

The exposure of enormous salt bottoms has resulted in devastating salt storms and, accompanied by dramatic climatic changes, has significantly changed the ecology of the entire region. Twenty of the 24 fish species native to the Aral Sea have disappeared and the annual fish catch, which employed 60,000 people in the 1950s, has been reduced to zero (Micklin 1988, p. 1172). The plight of the city of Muynak is typical: in the 1960s it was a thriving sea port and fishing center; now it is land locked and more than 40 miles from the Aral Sea (Hofheinz 1992). The Soviet government even shipped frozen fish 1,750 miles in an attempt to keep the processing plant open (Ellis 1990, p. 81).

It is evident that the ecological disasters of the Soviet Union and Eastern Europe are mind-boggling. Philip P. Micklin, a geography professor at Western Michigan University and a leading authority on the Aral Sea said, "I doubt if there has ever been an environmental problem of this magnitude" (Ellis 1990, p. 76). Another author has called Eastern Europe "the dirtiest, most degraded region on Earth" (Solomon 1990, p. A14).

But of course the question that is most pressing is why did the centrally planned economies do so poorly in providing for the environment? Why did industrialization and modernization bring so much more havoc than in other economies? In answering that question it is useful to review the development of modern economic theory as it applies to externalities.

Externalities in Modern Economic Theory

Arthur C. Pigou is generally credited with being the first to formalize the concept of an externality. Writing in 1912, he discussed how a factory would produce more than the optimal amount of smoke because of a divergence between private and social costs.² He expanded on this discussion in his *The Economics of Welfare* ([1920]1962, pp. 185–6), where he wrote:

Corresponding to the above investments in which marginal private net product falls short of marginal social net product, there are a number of others, in which, owing to the technical difficulty of enforcing compensation for incidental disservices, marginal private net product is greater than marginal social net product. Thus, incidental uncharged disservices are rendered to third parties when the game-preserving activities of one occupier involve the overrunning of a neighboring occupier's land by rabbits—unless, indeed the two occupiers stand in the relation of landlord and tenant, so

²Pigou, *Wealth and Welfare* (1912).

that compensation is given in an adjustment of the rent. They are rendered, again, when the owner of a site builds a factory there and so destroys a great part of the amenities of the neighboring sites; or, in a less degree, when he uses his site in such a way as to spoil the lighting of the houses opposite; or when he invests resources in erecting buildings in a crowded centre, which, by contracting the air space and the playing-room of the neighborhood, tend to injure the health and efficiency of the families living there.

Pigou's analysis has become so much a part of standard economic analysis that it is taught in almost all introductory classes. The concept of a divergence between social and private cost is also the basis of much public policy work, especially in environmental areas.

However, Ronald Coase (1960) enriched and corrected the Pigouvian framework by pointing out that in a world of well defined property rights and zero transaction costs exchange will eliminate externalities and insure that resources are allocated to their highest valued use. Thus, to use Pigou's example, if the game-preserver wanted to take actions that increased the rabbit population on a neighbor's land he could purchase the right to do so, if that right was held by the neighbor. On the other hand, if the game-preserver held the right to take actions that increased the rabbit population, the neighbor could compensate the game-preserver in order to prevent such an increase.

In such a world the actions that resulted would be socially optimal. If the game-preserver valued the increase in game more than the cost to the neighbor of the nuisance of more rabbits, he would either purchase the right to do so from the neighbor, or would not sell the right to prevent such an increase to the neighbor. On the other hand, if the neighbor found the cost of the rabbits to be greater than the value to the game-preserver of the increase in game, he would, depending on the initial allocation of rights, either purchase the right to be rabbit free, or not sell the right to increase the number of rabbits.

Note that Pigou understands the possibility of compensation in his description of the landlord tenant relationship, but fails to comprehend how property rights would produce the same situation even without a preexisting contract.

The Coasian approach shifts the externality discussion in that it removes us from the institutional vacuum that Pigou seems to operate in and thus causes one to focus on property rights and transaction costs. For public policy purposes it becomes important to see that rights are defined and that institutional barriers do not make transactions costs unduly high.

A final important step in the analysis of externalities came with the recognition that property rights are not static, but develop in response to economic conditions.³ As resources become more valuable it behooves people to expand effort on definition and enforcement activities. Such activity, when combined with Coasian exchange, can be an important element in the reduction of externalities.

For instance, in the United States when settlers moved West where water was scarce, the doctrine of prior appropriation evolved and replaced the riparian system that had dominated in the East. Under the riparian doctrine all owners of stream banks had co-equal rights to undiminished flows. Such a system made no allowance for diversion of water and had no provision for trading of rights. However, in the West diversion was an important use, but removing water from the stream reduced the water available for downstream users. Thus, under the riparian system the upstream diverter had no incentive to calculate the full costs of his actions and he imposed a negative externality on those downstream.

In response to this externality and to the increased value of water, a different system of rights developed. Under the new doctrine of prior appropriation, the first appropriator was granted exclusive right to the water and later appropriators' rights were conditioned on the prior rights of those who had gone before. Diversion was also permitted, and the exchange of water rights was allowed. Thus, a significant externality was eliminated.

Even in cases where high transaction costs prevent trades, efficient solutions of conflicts can emerge nevertheless. According to Bruce Yandle (1991, pp. 5–6):

Common Law rules of liability, which can be traded using the vehicle of mutually agreeable contracts, and legal processes that provide damage payments may substitute for the actual trade of environmental rights in a world with positive transactions costs.

Thus, allowing common law liability rules to evolve has been a historically important way of solving externality problems. For instance, under such a system English Common Law assigned a right of liability to downstream users in the case of water pollution and thus required payment of damages when such pollution occurred (Yandle 1991, p. 16).

Externalities and the Centrally Planned Economy

The above synopsis of the development of externality and property rights theory yields useful insights into environmental problems

³For elaboration of this argument, see Demsetz (1967, pp. 347–49); Anderson and Hill (1975, pp. 163–79); and Umbeck (1981).

under central planning. Pigou's discussion of the divergence between private and social net product is particularly applicable and is more general than usually imagined. Although Pigou is mainly cited for his insights concerning negative spillovers such as smoke, it is clear that he is discussing any situation where a decisionmaker acts without taking full account of all costs and benefits. But such a description fits well the manager and the central planner in command economies. Not just environmental costs, but many other costs and benefits are ignored by the decisionmaker in these societies. In fact, one could describe such an institutional framework as one where Pigouvian externalities abound in every facet of life. However, since environmental problems are the focus of this paper it is useful to examine more closely central planning to see just how these types of externalities come about.

Incentive Problems

Much is external to those responsible for making decisions under socialism for two reasons: incentive problems and inadequate information. Incentives to act correctly are missing because of a lack of private property rights. Without property rights, it is virtually impossible to hold those who control resources accountable since they have no claim on changes in the value of the resource. Even when they ignore whole sets of costs their own welfare may be little altered.

For instance, it is very easy to allow political considerations to dominate economic choices, often to the detriment of the environment. In many cases, the overriding political imperative in socialist economies was increased industrial output, with little attention paid to accompanying air or water pollution (Goldman 1972, p. 64).⁴ Since property rights are a mechanism for appropriately assigning costs, their absence meant planners could opt for increased physical output without having to compensate those who bore costs in the form of pollution.

Krakow, Poland is a prominent example for a decisionmaking process in which the crassest of political reasons were the guiding principles. This city is the location for a large steelwork whose effluent is

⁴One could also argue that preferences were such in these economies that the choices that were made were optimal; people simply preferred higher levels of pollution because they wanted the jobs and incomes that accompanied them. However, the fact that environmental concerns were at the heart of many of the first democracy movements in Eastern Europe would lead one to doubt this conclusion. Also, pollution levels in Eastern Europe are much higher at similar levels of industrialization and income than in Western Europe. It is not clear why preferences would be so radically different between the two regions.

doing enormous damage to the historic buildings and sculptures of the city. A journalist, Jon Thompson (1991, p. 44) reported:

The next day, I met Stanislaw Juchnowicz, a distinguished-looking architect with spectacles and silver hair, president of the Polish Ecology Club. "Why," I asked him, "was a huge steelwork built next to a city of such historic importance?"

He paused, weighing his answer. "You must understand, it was a political decision. There is no iron ore here, and we had very little industry. According to the theory of our communist masters, the wage-earning class was supposed to have a leading role in society. In the 1950's all the countries under communist rule underwent massive industrialization. Krakow was a university town with very few wage earners. Putting the steelworks here was a deliberate attempt to destroy the old order by creating a class of wage earners where none existed before."

In a system of private property rights, residual claimancy provides a means whereby decisionmakers cannot be oblivious of costs in taking actions like those described above. If one had to purchase all resources in the market place, it would be difficult to justify placing a steel plant in a particular location simply to alter the social structure of a city.

The general inefficiency of production under socialism is another indicator of the lack of incentives to prevent waste. Czechoslovakia consumes about three times the energy of the average western nation per unit of output (Bingham 1991, p. 101). In the former Soviet Republics manufacturing uses four times as much energy per unit of GNP as in the United States (Hofheinz 1992, p. 110). Chemical plants in the Soviet Union for many years emitted large amounts of a potent pollutant, fluorine, into the atmosphere (Komarov 1980, p. 92). Despite numerous studies by engineers that showed that the fluorine could be recovered at a profit and sold to other enterprises, the plant managers found it easier to continue to pollute. There was no effective system in place whereby a manager was rewarded for taking such cost reducing and environment improving actions.

Private rights are also a way of ensuring that those who control resources take account of the effect of present actions on future resource values. Since the price of a resource reflects the capitalized value of the future income stream from that resource the decision makers have every reason to take account of future effects of their actions.

However, under a system of central planning and state ownership, it is much more likely that the resource manager will be short-sighted. The lack of property rights means the decisionmakers are not penalized through decreases in the asset value when they ignore

future consequences of their actions. Boris Komarov (1980) provides the following examples for the Soviet Union:

A. P. Kitaev, director of the Pakhvistnevski Logging Enterprise in Iaroslav Province, admits in passing that in the ten years he has worked there, the management has never questioned him about fulfillment of the plan for replanting seedlings, but only about the plan for fellings [p. 100].

One sorry tale was the breeding of polar foxes on farms. Sorry, because to feed the voracious foxes, whales were killed, including, of course, very rare species, females, and the young. Harpooners easily distinguish authorized from prohibited species by their spout; but they are always paid for the number of carcasses, and there has never been a case in which a harpooner was punished for a mistake. They *are* punished for nonfulfillment of the plan [p. 123].

Finally, the lack of private property rights meant the legal system was ineffective in terms of stopping pollution. One of the features of private property is the ability to stop other people from taking actions that damage your property. The courts provide a means of redress when such actions are threatened or take place. However, under socialism the lack of private rights meant individuals could not use the system to prevent harm to property. Farmers in Bulgaria were well aware that heavy metals carried in irrigation water were lowering crop yields and that the sources of the pollution were readily identifiable. However, with no property rights the farmers had no actionable claim against those causing the problem.

Likewise, despite numerous laws designed to prevent pollution, factory managers had little incentive to install pollution control measures. In certain cases fines were levied against the factory, but the manager was allowed to include these fines in his budget allocation from the central ministry. The fact that power and law were in the same hands meant that there was little chance of redress through the legal system. One part of the state would have been suing another arm of the state, with that same state deciding the case.

It is also important to note that the incentive system under socialism discouraged diversity. People who had divergent preferences or opinions about resource use found it difficult to act on their desires. Minority viewpoints are seldom recognized under central planning and this was certainly the case in the area of ecological concerns. Under a system of private rights, an individual who believes a resource is being over-exploited has only to convince the resource owner to sell it; then that individual can preserve it as he wishes. However, no such allowance for divergent preferences exists when resources are under political control. It is essentially a "winner takes

all" system. Since the prevailing sentiment of those in power in Communist regimes favored forced industrialization and the rapid use of resources, that position dominated all decisions.

Information Problems

The argument that socialism suffers from a badly structured incentive system is a powerful one. However, environmental problems resulted from bad information as well as from poor incentives. Even if resource managers would have wanted to behave in an ecologically responsible manner they would have found it difficult to do so because of the low quality of information they had to work with.

Prices represent an efficient network for consolidating information and providing rapid feedback to decisionmakers. However, such an information flow depends upon private property rights, because prices represent bids and offers to take actions with regard to property. Only when the decisionmaker faces the true opportunity costs of his actions do prices reflect economic reality. Since property rights are not fully defined and enforced in most market economies, prices do not always convey completely accurate information. Pigouvian externalities do exist. However, such occurrences represent aberrations, and most of the time people act on reasonably good information under these systems. However, under central planning and with state ownership of property exactly the opposite is the case; prices that accurately reflect reality are aberrations. Also, there is no feedback mechanism by which prices are adjusted so they become more accurate. In the face of such an information blackout, it is not surprising that environmentally unsound decisions dominate.

The difficulty of securing accurate information is illustrated by two examples from Bulgaria. In the winter of 1991 I visited a large shipyard in Varna. When questioned about the possibilities for privatizing the facilities the manager assured me his operation would be easy to sell off because it consistently showed a profit. However, in a private conversation, another official pointed out that the shipyard was allocated its steel at a price that was one third of the world price. Thus, the manager had no sound basis for making his judgments about profitability.

In a similar vein, Bulgaria has long been a leading producer of forklifts for the world market. Seen as one of the successes in a sea of failures, its forklift factory is an important source of foreign earnings and national pride. Lead is an important material in forklift production and is produced in the town of Kardszhali. However, the smelter is also causing massive health problems to area residents. By U.S. standards all of the children of the town have dangerously

high concentrations of lead in their blood. Because of the importance of the lead manufacture to the forklift industry, however, officials have been reluctant to close down the lead and zinc complex.

It would seem that the traditional conflict between jobs, economic growth, and environmental quality is at play here. However, recent evidence indicates that the Ministry of Industry supplies lead and zinc to the forklift industry at one quarter of world prices (Harden 1990). Thus, it may well be that no conflict exists. An industry that is heavily subsidized and a net loser for the economy is also responsible for enormous environmental costs.

The lack of accurate prices is also reflected in other ways. There is no differential rent system in most communist countries; land and timber close to population centers are priced the same as more distant resources, despite the fact that the closer land has a higher amenity value. Therefore, in making their calculations of which resources to exploit first, planners have no incentive to use the more distant ones. As a consequence, forests in the vicinity of major population centers are rapidly harvested, while others further removed are left unharvested (Goldman 1972, p. 169).

Not only is information of poor quality under socialism, there is also little mechanism for improving it. Resource managers are seldom rewarded for generating good information nor are they punished for conveying poor information. Therefore mistakes are slow to be corrected. Komarov (1980) reports that a state farm close to the Crab Lakes in the Soviet Union was mandated to fulfill a reclamation plan. Several of the lakes were drained, but according to Komarov (pp. 51-52):

The draining of silty shallows added not a single hectare of new meadowland to the state farm. It was known beforehand that this would be the case, and the farm probably didn't even need the land. In what other economy would it be possible to squander thousands of rubles digging drainage canals and destroying lakes when it was known beforehand that no one would benefit at all from it?

Although prices are a crucial source of information, they are not the only one that is important to maintaining ecological integrity. The role of the press can be instrumental in making people aware of major environmental disruptions and identifying those responsible. However, a free press depends on private property rights. It is the opportunity to purchase newsprint, rent production facilities and hire reporters without state approval that gives the press its freedom. The fact that all property was state owned in socialist economies

meant the press played no role in monitoring the environment and holding responsible those who despoiled it.

When the Chernobyl accident occurred in the Soviet Union and the nuclear cloud was headed for Bulgaria the top government officials secreted themselves away in special shelters with ample supplies of food and water. Bulgarian citizens, however were not even informed that the disruption had happened and continued to eat contaminated fruit and vegetables.

In Czechoslovakia the government had data concerning the amount of sulfur dioxide in the air in different regions but kept it secret (Dickman 1990). Scientists in East Germany also generated detailed information on pollution levels and types but that was also not made available to the general population (Charles 1990). In Bulgaria it was a crime to publish scientific findings about industrial pollution (Harden 1990).

Finally, information was significantly biased because of ideological considerations. The labor theory of value, a major component of Marxist ideology, holds that cost is determined by the amount of labor invested in producing an item. Since natural resources exist without any such investment it would violate Marxist ideology to charge a price for water, land, or minerals. Thus, it is not surprising that overuse of these resources occurs. When the basic water law was passed for the Soviet Union in 1970 a leading jurist said:

The principle of using water resources permanently and without charge is inherent in the law of water use in the USSR. Natural reserves of water are always allotted for use without charge and in the overwhelming majority of cases for an indefinite period of time [Goldman 1972, p. 40].

Lack of Exchange Solutions

I have argued above that Pigouvian externalities abound in socialist economies and that they are responsible for much of the environmental disruption that has occurred. But, as Coase (1960) has pointed out, many potential externalities are resolved through exchange. However, not only does socialism ensure that externalities will occur, it also puts significant barriers in the way of their resolution through exchange. Of course, the fact that property rights are not allowed to formally exist means exchange is limited. Nevertheless, quasi-rights do evolve, and they could be exchanged if the system allowed.

One author has suggested that in the Soviet Union "ruling class control constitutes a set of de facto property rights over the goods of the economy" (Schap 1988, p. 393). Those property rights exist because managers and ministry officials have decisionmaking authority.

However, central planning mitigates against exchange relationships among ministries or production units. Therefore, the potential for externalities to be resolved through contracts is lessened because of high transaction costs.

For instance, in the Georgian Republic in the Soviet Union, contractors avail themselves of large amounts of sand from the Black Sea beaches. These make for a cheap source of materials for concrete, and, because the beaches are state property, no one can formally object. However, such actions have had a serious impact on the value of waterfront property. As early as 1960 it was estimated that Soviet Black Sea beaches had shrunk by 50 percent, and resort hotels have collapsed (Goldman 1972, pp. 158–59). The ministry in charge of resorts has, of course, had the value of its quasi-property rights significantly diminished. However, even if the loss to that ministry is greater than the gain to the ministry that controls construction, and hence is responsible for using up the beaches, there is no formal mechanism by which the appropriate officials can enter into a mutually profitable exchange.

A similar situation has occurred with the building of dams for hydroelectric projects and the diversion of water for irrigation. In many cases these projects have had a serious impact on fishing or on farmland.⁵ Again, *de facto* property rights do exist in many of these situations, especially on the state and cooperative farms. The farm workers and managers control the resources to a certain extent and have some degree of residual claimancy. However, even when the potential loss to the farm or to the fishery far exceeds the gains from the water project, no exchange is possible that would allow the resolution of the conflict. Neither can the farms (or the Ministry of Agriculture) compensate the water developers in order to stop the projects, nor can they demand compensation if the projects go forward.

A similar situation occurs when farm land suffers from pollution. In Bulgaria over 70 percent of the land suitable for farming has been affected by industrial pollution, much of it carried by irrigation water (Harden 1990). Again, those who operate the cooperative farms are very much aware that their property has been dealt a severe blow. However, high transaction costs keep them from entering into an exchange to resolve the problem.

⁵See, for instance, Ellis (1990), Micklin (1988), Goldman (1972, p. 65), and Komarov (1980, p. 57).

No Evolution of Property Rights

The final area which causes environmental problems under socialism is the lack of a mechanism by which property rights can evolve. As countries industrialize certain resources become more scarce, particularly clean air and water. Since defining and enforcing property rights is very much an economic activity, it does not necessarily pay to invest in such activity when resources are not scarce. However, with increasing scarcity individuals will find it worthwhile to make such investments and property rights will develop.⁶

Centrally planned economies have no mechanism for such development and state ownership of property usually means that a rigid system of rights exists. Hence, even though certain resources become more valuable over time no evolution of rights occurs. In the case of the Black Sea beaches discussed above, it is not clear that formal property rights existed to all of the waterfront prior to the Communist revolution of 1917. However, in an economy where entrepreneurial initiative and freedom of contract were allowed, one can be assured that rights to that property would have developed as beaches became more valuable.

Likewise, in the case of the Aral Sea fishing rights were largely a matter of historical precedent. However, as those rights became more valuable, pressure would have led to formalization and codification. Under an evolutionary legal structure like the Common Law, actionable claims would have come into being and would have offered the Aral fishermen either legal redress or the opportunity for market exchange when massive irrigation projects threatened their livelihood.

Conclusion

It is not surprising that socialist economies have experienced significant environmental problems. The structure of incentives and information under socialism ensures the existence of a multitude of classical Pigouvian externalities. The lack of well-defined property rights and the existence of artificially high transaction costs mean that Coasian exchange cannot function to resolve conflicts when the use of a resource affects the value of another resource. Socialist economies also have rigid systems that do not allow property rights to develop in response to changing values. All of these elements mean that ecological disruption is likely, and the record amply testifies to the truth of that conclusion.

⁶For a more complete discussion of this concept as applied to environmental problems, see Anderson and Leal (1991).

References

- Anderson, Terry L., and Hill, Peter J. "The Evolution of Property Rights: A Study of the American West." *Journal of Law and Economics* 18 (April 1975): 163-79.
- Anderson, Terry L., and Leal, Donald R. *Free Market Environmentalism*. San Francisco: Westview Press, 1991.
- Bingham, Sam. "Czechoslovakian Landscapes." *Audubon* (January 1991): 92-102.
- Charles, Daniel. "East German Environment Comes into the Light." *Science*, 19 January 1990, pp. 274-76.
- Coase, Ronald H. "The Problem of Social Cost." *Journal of Law and Economics* 3 (October 1960): 1-44.
- Demsetz, Harold. "Toward a Theory of Property Rights." *American Economic Review* 57 (May 1967): 347-59.
- Dickman, Steven. "Pollution as Czech Public Enemy Number One." *Nature* 344, 8 March 1990, p. 91.
- Ellis, William S. "The Aral: A Soviet Sea Lies Dying." *National Geographic* (February 1990): 73-92.
- Feshbach, Murray, and Friendly, Alfred Jr. *Ecocide in the USSR*. New York: Basic Books, 1992.
- Goldman, Marshall I. *The Spoils of Progress: Environmental Pollution in the Soviet Union*. Cambridge, Mass.: MIT Press, 1972.
- Harden, Blaine. "Key Bulgarian Industry Is Poisoning Children." *The Washington Post*, 13 August 1990, pp. A1, A14.
- Hofheinz, Paul. "The New Soviet Threat: Pollution." *Fortune*, 27 July 1992, pp. 110-14.
- Komarov, Boris. *The Destruction of Nature in the Soviet Union*. White Plains, N.Y.: M.E. Sharpe, 1980.
- Lange, Oskar, and Taylor, Fred M. *On The Economic Theory of Socialism*. Minneapolis: University of Minnesota Press, 1938.
- Maremont, Mark; Kapstein, Jonathan; and Schares, Gail E. "Eastern Europe's Big Cleanup." *Business Week*, 19 March 1990, pp. 114-15.
- Micklin, Philip P. "Desiccation of the Aral Sea: A Water Management Disaster in the Soviet Union." *Science* 241, 2 September 1988, pp. 1170-76.
- Painton, Frederick. "Where the Sky Stays Dark." *Time*, 28 May 1990, pp. 40-42.
- Pigou, Arthur C. *Wealth and Welfare*. London: Macmillan, 1912.
- Pigou, Arthur C. *The Economics of Welfare*. 1920. London: Macmillan, 1962.
- Schap, David. "Property Rights and Decision Making in the Soviet Union: Interpreting Soviet Environmental History." *Economic Inquiry* 26 (July 1988): 389-401.
- Simons, Marlise. "Eastern Europe: The Polluted Lands." *New York Times Magazine*, 29 April 1990, pp. 30-35.
- Solomon, Laurence. "The Best Earth Day Present: Freedom." *Wall Street Journal*, 20 April 1990, p. A14.
- Thompson, Dick. "The Greening of the U.S.S.R." *Time*, 2 January 1989, pp. 68-69.
- Thompson, Jon. "East Europe's Dark Dawn." *National Geographic* (June 1991): 36-63.

ENVIRONMENTAL PROBLEMS

- Umbeck, John. *A Theory of Property Rights With Applications to The California Goldrush*. Ames: Iowa State University Press, 1981.
- Yandle, Bruce. "Escaping Environmental Feudalism." Paper prepared for a conference on "Free Market Environmentalism: The Role of the Market in Environmental Protection," Lewis and Clark College's Northwestern School of Law, 13 April 1991.

THE POLITICAL ECONOMY OF MANDATED SPENDING

Marilyn R. Flowers

Introduction

It is widely acknowledged that government spending and taxation are not the sole sources of government influence on resource allocation and the distribution of income. Modern governments also wield vast regulatory powers and, in many circumstances, direct regulation of private sector entities is an alternative to government taxing or spending. Richard Wagner, for example, asserts that “a central principle of public finance is that any statute or regulation can be translated into a budgetary equivalent” (Wagner 1989, p. 108). Harvey Rosen, in his best-selling public finance text, refers to the “‘hidden’ costs of government,” in pointing out that many regulations have important and large effects even though they require minimal government outlays (Rosen 1988, p. 17).

The failure of the budget document to reflect the economic impact of regulation is potentially more than a pure measurement issue. There is a perennial debate in the United States over the desirability and feasibility of imposing constitutional limits on the activities of the public sector, particularly at the national level. To the extent that the most commonly proposed constraints, spending limits, and balanced budget requirements are expressed in terms of fiscal variables alone, concern has been expressed that the availability of regulation will provide a means for the political process to circumvent those constraints. At a 1980 conference, for example, no fewer than three of the participants expressed concern that increased regulation would provide an avenue for the political process to subvert any constitutional limits on its spending or borrowing authority (Moore and Penner 1980, pp. 48, 54, and 60). Similarly, Aaron Wildavsky

raised the issue of regulation as an "end run" around a constitutionally imposed spending limit.¹

This paper takes a critical look at the public choice aspects of the choice between direct government spending and government mandated private expenditures.² Some similar issues arise with respect to intergovernmental relations when one level of government imposes mandates on the spending levels and patterns of a jurisdiction or jurisdictions below it in the federal hierarchy. These issues will not be considered in this paper. Under what circumstances is the government likely to require that private sector entities, either firms or consumers, spend their own after tax resources to achieve some publicly (i.e. politically) determined objective and what are the implications of this choice for imposing constitutional fiscal constraints? The following propositions will be developed.

1. A necessary condition for government intervention is the existence of some constituency that favors it. However, supporters of intervention may have preferences regarding the form of such intervention and only in a limited number of cases will direct government spending and mandated private spending be regarded as substitutes in the achievement of the constituents' goals. Rarely, if ever, will the two forms of intervention be regarded as perfect substitutes. Fiscal limits will not influence these preferences.
2. A balanced budget amendment may have little impact on public sector size in general, and probably does not, in and of itself, create a stronger constituency in favor of mandated private spending.

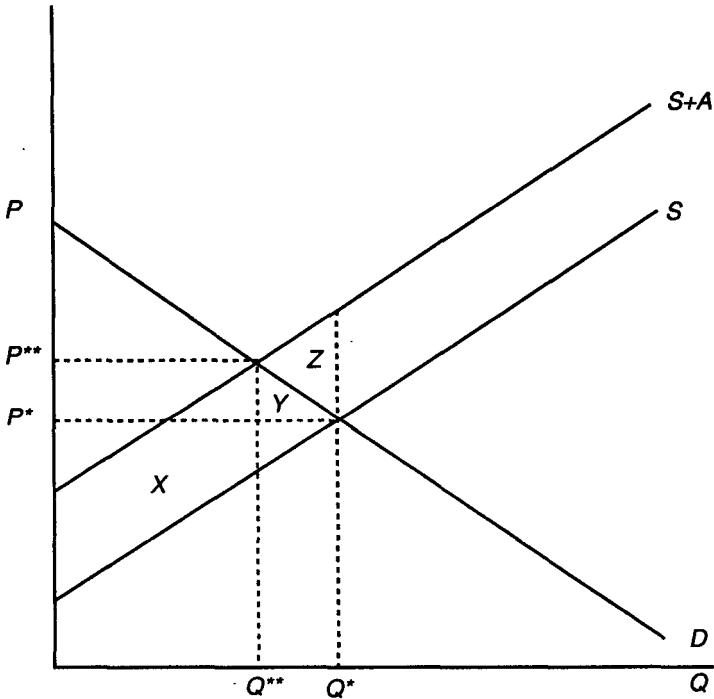
The Demand for Intervention

Much of the discussion of propositions (1) and (2) above will be based on the simple example presented in Figure 1. That figure depicts the market for some good produced by a competitive, industry. The market is initially in equilibrium with output Q^* and price P^* . Suppose that the government is considering a policy of requiring some safety provision in the industry. This provision will raise marginal cost by some amount, A . If the government mandates that the

¹Wildavsky argued that the substitution of regulation for direct spending was not a likely outcome. His reasons were very different from the ones examined in this paper, however.

²Some similar issues arise with respect to intergovernmental relations when one level of government imposes mandates on the spending levels and patterns of a jurisdiction or jurisdictions below it in the federal hierarchy. These issues will not be considered in this paper.

FIGURE 1
MARKET RESPONSE TO MANDATES



industry pay for the safety provision, the industry supply curve will shift up by A to the curve labeled $S + A$, and equilibrium price and output in the industry will rise and decline to the levels P^{**} and Q^{**} , respectively. Note that the cost of the safety provision is the sum of the area labeled X , representing the resource costs of implementing the provision, and the triangle Y that represents additional losses in consumer and producer surplus.

An alternative policy option would have the government bear the cost of the safety provision out of public funds. The government could either purchase the safety device or, more plausibly, require that firms buy it, but provide a subsidy to cover the cost of purchase. This policy has very different implications for price and output in the industry. Because cost is unaffected at the firm level, there is no increase in price and corresponding decline in industry output.³

³There is the possibility that the extra taxes may somewhat reduce the demand for the good. This effect seems sufficiently small to be safely ignored.

Because the safety provision now affects Q^* units of output, the total cost of the provision is the area X plus the two triangles labeled Y and Z . The cost of the safety provision in the direct government expenditure case exceeds cost in the mandated private expenditure case by the area of triangle Z .⁴

The example thus far is contrived in that no source of demand for the safety provision is identified. It is difficult to understand why the government would be involved at all under such circumstances. The next step in the analysis is to consider some alternative sources of demand for the safety provision and to ask whether the potential demanders would have any preference between direct government expenditures and mandated private expenditures. In the discussion that follows, three examples will be considered. In each of the examples, the constituency with a demand for intervention has a distinct preference for the *nature* of the intervention. In only one of the examples do the supporters of intervention regard direct government spending and mandated private spending as substitutes. Even in that case, the two policy options are imperfect substitutes.

Two assumptions should be made explicit before proceeding with development of the specific models below. First, it will be assumed that consumers and producers are aware of the differential implications for market equilibrium of direct government spending and mandated private spending. Second, although some subset of taxpayers must constitute the potential demanders for intervention, for purposes of simplification, a general distinction will be made between taxpayers at large and individuals with a positive preference for intervention. Thus, taxpayers can be characterized as opposed to intervention via the direct government expenditure route and neutral with respect to the choice between no intervention and intervention via mandated expenditures. Any constituency with a demand for direct spending, on the other hand, can be simply characterized as a group for whom the benefits of the spending exceed their own share of the tax costs.

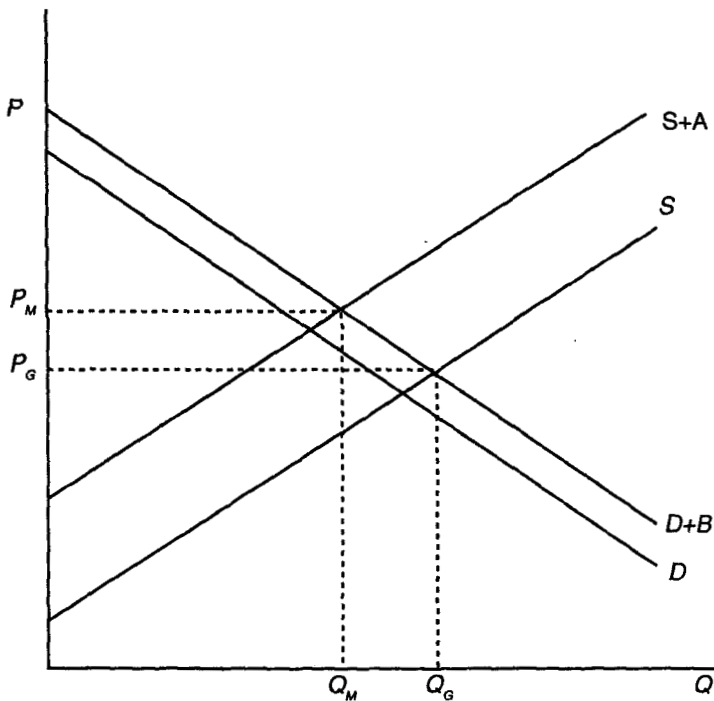
Example 1: Demand for Direct Government Spending

The first example to be considered is one in which the safety provision is of some value to consumers but not of sufficient value to lead to market provision. Suppose consumers place a value of B

⁴In addition to the added cost represented by triangle Z , the government spending option also should include the excess burden associated with the marginal taxation required to finance that option. This point is made by Summers (1989) in his comparison of the relative efficiency of mandates and direct government spending.

per unit on the safety provision. Referring to Figure 2, the presence of the safety provision shifts the demand curve up vertically by that amount. The safety provision increases marginal cost by A as was the case in the original example. If $B > A$, the market process will provide the safety without government intervention. If, on the other hand, $B < A$ as depicted in Figure 1, the safety provision will only be provided if the government intervenes.

FIGURE 2
DEMAND FOR DIRECT SPENDING



If the safety provision is financed by direct government spending, there will be an increase in both consumer and producer surplus at the new market equilibrium price and quantity of P_G and Q_G respectively. Therefore, both consumer and producer interests should have a demand for intervention via the direct government spending route. However, the situation is very different if the safety provision is the result of mandated expenditure by the industry. Market equilibrium in this case is characterized by P_M and Q_M . Both producer and consumer surplus are reduced by this form of intervention and, correspondingly, there is no political demand for interven-

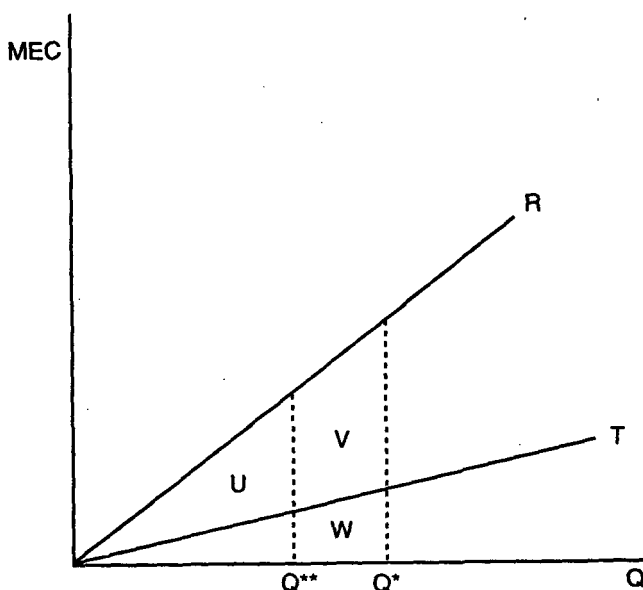
tion of this sort. The two possible political outcomes thus should be no intervention or government intervention via the direct expenditure route with the actual choice reflecting the political balance between the interests of consumers and producers with relatively small total but large per capita benefits and the opposition of taxpayers who have large total but small per capita losses.

Example II: Substitutability between Mandated and Direct Spending

A scenario in which mandated private expenditures acquire some political plausibility exists if the good in question is associated with some externality or third party effects. Suppose, for example, that production of the good generates some risk of harm to third parties who are neither producers nor consumers of the product. In Figure 3, the marginal external cost of that risk as a function of industry size is depicted by the function labeled R if the safety device is not used and by the function labeled T if the safety device is used. Industry

FIGURE 3

SUBSTITUTABILITY BETWEEN MANDATES AND SPENDING



size in the absence of government intervention is Q^* as before. If the safety device is implemented as a result of direct government expenditures, cost savings to third parties are represented by the

sum of the areas of the triangle labeled U and the quadrangle labeled V . If, on the other hand, the safety device is implemented as a result of mandated industry expenditure with a corresponding shrinkage of the industry to Q^{**} , cost savings to third parties are increased by the area of the quadrangle labeled W .

In this case, the demanders of intervention have a preference for mandated private spending, but will garner some smaller net gain if the direct government spending policy is adopted. The political choice of the manner of government intervention, assuming it occurs, then should reflect a balance between the greater demand for mandates and the opposition from taxpayers to direct spending and a coalition of producer and consumer interests who prefer direct spending to mandated private spending.

Example III: Demand for Mandated Private Spending

Although regulation is often characterized as a means of imposing the cost of achieving public objectives on the affected industries and their resource suppliers, particularly labor, rather than on the taxpayers at large, various scholars have noted that the opposition from these groups is less vigorous than might be expected and in some cases is almost nonexistent. It is hypothesized that in some cases the regulation is welcomed because it enables a group of firms within an industry or perhaps an entire industry to gain a competitive advantage over another group of firms or another industry.⁵ Suppose, for example, that the cost of a worker safety provision was smaller for one group of firms in an industry than for another group. If government policy were to require that all firms provide the added safety, product price must rise to cover the cost of that provision to the marginal (high cost) firms in the industry. This creates inframarginal rents to other firms who are able to satisfy the government requirement at lower cost. This latter group of firms would benefit from the regulatory provision.

In context of the discussion in this paper, the situation described above creates a constituency that would support mandates and oppose direct spending. Simultaneously, an opposing constituency exists that would support direct spending and oppose mandates. A simple example can illustrate how mandated spending can create rents for one group of firms in an industry while direct spending creates rents for another group of firms. Suppose that the industry in question is comprised of two types of firms. Both types hire in a

⁵See, for example, Ackerman and Hassler (1981), Bartel and Lacy (1985, 1987), Fuess and Lowenstein (1990), Maloney and McCormick (1982).

competitive national labor market, but there is a risk to workers which is unique to the industry. The risk can be eliminated at a cost of CI per worker in type I firms and CII per worker in type II firms. Workers place a value of A on elimination of the risk. Assume that $CII > A > CI$. This implies type I firms will eliminate the risk even in the absence of government intervention, but that type II firms will not. Given a national wage of W , employees of type I firms will receive that wage, and the labor cost per worker to those firms will be $W + CI$. Employees of type II firms will receive a wage of $W + A$. In the absence of government intervention, the market price of the product will adjust to a level consistent with positive economic rent for type I firms reflecting their lower labor costs and zero rent for type II firms.

If the government mandates that all firms provide a safe workplace, the rent earned by type I firms will increase. The mandate will have no effect on labor costs of those firms, but labor costs will increase, to $W + CII$ per worker, for type II firms. This increase in marginal cost of production will lead to an increase in product price and a corresponding increase in the economic rent accruing to type I firms. In contrast, type II firms will gain and type I firms lose if the government finances the safety provision directly. Relative to the no-intervention equilibrium, government provision will cause labor costs to fall by CI for type I firms and by A for type II firms. Type II firms will now be the low cost firms and correspondingly recipients of positive economic rent.

The preceding example is illustrative of a situation in which there are competing constituencies for mandates and direct public spending. This is in contrast to example I, in which there is a constituency only for direct public spending, and example II in which a single constituency favoring government action regards mandates and direct spending as gross substitutes. In example III, neither of the competing constituencies regards mandates and direct spending as substitutes. Each will prefer one form of intervention but oppose the other.

Constitutional Constraints and Mandated Spending

As noted in the introduction, concern has been expressed that attempts to limit the size and influence of the public sector via imposition of fiscal constraints may be of limited effectiveness if political actors are able to circumvent such rules by use of the government's regulatory authority. However, for this to occur, there must be a constituency that would support regulation as an alternative to

direct expenditure by the government. The preceding discussion suggests that the circumstances in which such a situation could occur are somewhat limited. In examples II and III above, such constituencies exist. However, two additional conditions must be met if increased regulation is to be a *consequence* of the imposition of fiscal constraints. First, it must be the case, prior to imposition of the constraints, that either the spending or the nonintervention option had been chosen in the relevant markets. The second condition is that imposition of fiscal constraints strengthen the political support for regulation relative to either of those options.

A weakening of the political influence of constituencies that favor government spending would occur if the imposition of fiscal constraints were to strengthen the influence of the taxpayers at large who are the primary opponents of the spending option.⁶ In the case of a balanced budget constraint, a presumed source of such a change might be stronger taxpayer opposition to government spending when that spending must be financed by taxation than would be the case if public debt were available as the marginal source of government finance. Do taxpayers regard deficit finance as a means of reducing the marginal cost to them of public spending? If public borrowing is a legitimate way of shifting burden onto future generations, it might appear that the answer to this question would be yes. However, when couched in a proper framework of the opportunity cost of government spending, this proposition can be challenged. Generally speaking, there are two ways in which the government might increase the size of the annual deficit. The first is to increase public spending; the second is to hold spending constant but reduce taxes. Viewed from this perspective, the representative taxpayer should evaluate the marginal cost of government spending on the basis of his or her marginal tax share as determined by the prevailing tax structure.

Public debt, in the absence of Ricardian behavior, is an intergenerational transfer of wealth. The beneficiaries of the transfer have the option to consume their added wealth in the form of an increase in either public or private consumption. The marginal opportunity cost of increased public spending remains reduced private spending. There remains a wealth effect, of course. Assuming that both public and private consumption are normal goods, imposition of a balanced budget rule can be predicted to reduce demand for both with a

⁶This model of the political process follows Becker, who, in a seminal paper, developed the argument that "increased influence by some groups decreases the influence of others by equal amounts" (Becker 1983, p. 376).

corresponding increase in domestic saving and investment. However, this does not suggest that proponents of the type of spending discussed here will be weakened relative to proponents of other government spending programs. Nor does it suggest a relative weakening of the political influence of the constituencies that oppose mandated expenditures.

Conclusion

The indisputable incompleteness of the government budget document in reflecting the economic cost of the public sector has made enforceability a key issue in discussions of the feasibility of imposing meaningful fiscal limits. The presumed justification for imposing such limits is that important ramifications of the aggregate outcome of the budgeting process are not adequately reflected in the piecemeal issue by issue process by which the individual components of the budget are determined. An overall fiscal constraint, by itself, does not alter underlying incentives to ignore aggregate implications, but rather is an attempt to arbitrarily limit the ability of the system to respond to those incentives. Because the incentives remain, the constraint will be ignored if its terms are unenforceable. However, concern with enforceability must also reflect a careful understanding of the nature of the incentives that lie behind particular hypothesized avenues of escape. The analysis in this paper suggests that one commonly expressed source of concern, direct regulation of private sector activity, may well be overstated.

References

- Ackerman, Bruce, and Hassler, William. *Clean Coal/ Dirty Air*. New Haven: Yale University Press, 1981.
- Bartel, Ann, and Thomas, Lacy Glen. "Direct and Indirect Effects of Regulation: A New Look at OSHA's Impact." *Journal of Law and Economics* 28 (1) (1985): 1-26.
- Bartel, Ann, and Thomas, Lacy Glen. "Predation through Regulation: The Wage and Profit Effects of the Occupational Safety and Health Administration and the Environmental Protection Agency." *Journal of Law and Economics* 30 (2) (1987): 239-64.
- Becker, Gary S. "A Theory of Competition Among Pressure Groups for Political Influence." *Quarterly Journal of Economics* 98 (3) (1983): 371-400.
- Fuess, Scott, and Lowenstein, Mark. "Further Analysis of the Theory of Economic Regulation: The Case of the 1969 Coal Mine Health and Safety Act." *Economic Inquiry* 28 (2) (1990): 354-89.
- Maloney, Michael, and McCormick, Robert. "A Positive Theory of Environmental Quality." *Journal of Law and Economics* 25 (1) (1982): 99-124.

- Moore, William S., and Rudolph Penner, eds. *The Constitution and the Budget*. Washington, D.C.: American Enterprise Institute, 1980.
- Rosen, Harvey S. *Public Finance*. 2d ed. Homewood, Ill.: Richard D. Irwin, 1988.
- Summers, Lawrence H. "Some Simple Economics of Mandated Benefits." *American Economic Review* 79 (2) (1989): 177–83.
- Wagner, Richard E. *To Promote the General Welfare*. San Francisco: The Pacific Institute, 1989.
- Wildavsky, Aaron. *How to Limit Government Spending*. Berkeley: University of California Press, 1980.

REFORMING THE BONNEVILLE POWER ADMINISTRATION

Kenneth W. Costello and David Haarmeyer

Introduction

The Bonneville Power Administration (BPA) is the power marketing agency for the Federal Columbia River Power System (FCRPS), which also consists of the Pacific Northwest generating facilities operated by the Army Corps of Engineers and nonpower-related projects of the Bureau of Reclamation. Taken together, the accounting value of the FCRPS's total assets as of the end of September 1991 was over \$15 billion. In 1991, the BPA marketed over \$2.2 billion worth of electricity from 30 federally built powerplants.¹

Through its extensive transmission network (Intertie), the BPA markets power to a region encompassing the states of Washington, Oregon, Idaho, western Montana, plus portions of neighboring states. The Intertie makes up nearly 80 percent of this region's high-voltage transmission, with over 14,700 circuit miles of lines and almost 400 substations. The replacement value of the system is estimated by the BPA to be about \$10 billion.² In addition to marketing power, the BPA also promotes nonelectricity-related objectives, which include conservation, irrigation, and fish and wildlife protection.

Cato Journal Vol. 12, No. 2. (Fall 1992). Copyright © Cato Institute. All rights reserved.

Kenneth W. Costello is an Economist at Ohio State University where he conducts research on public utility regulation, and David Haarmeyer is a Policy Analyst at the Reason Foundation.

The authors wish to thank Jim Clarkson, Richard L. Gordon, Douglas Houston, James Johnston, Philip O'Connor, Lynn Scarlett, David Shapiro, Richard Stroup, and John Wenders for helpful comments. Responsibility for any remaining errors, of course, lies with the authors.

¹U.S. Department of Energy (1992).

²U.S. Department of Energy (1991a, p. 10).

Along with the four other federal power administrations,³ the BPA is required to give preference in the sale of power to utilities owned by public entities such as municipalities, public utility districts (PUDs), and rural electric cooperatives. In addition to also providing nonfirm power to investor-owned utilities, the BPA sells power directly to aluminum producers in the Northwest.

The BPA was established by Congress in 1937 to market and transport power from the Bonneville Dam on the Columbia River.⁴ Like the four other federal power marketing administrations, the BPA's original objective was to promote economic development by providing low-cost power, particularly to rural areas.⁵ At the time, it was believed that the highly concentrated privately owned electric industry was either unable, unwilling, or could not be trusted to provide power at reasonable prices.⁶

The U.S. electric industry has since undergone dramatic change. As a result of technological, economic, and regulatory changes, the industry's early monopoly structure has increasingly given way to a more open and competitive structure.⁷ The growing significance of independent suppliers of electricity in addition to the gradual interconnection of local utilities means that competition within the industry is likely to accelerate. By ensuring that prices more closely reflect

³The four are the Alaska Power Administration, the Southeastern Power Administration, the Southwestern Power Administration, and the Western Area Power Administration.

⁴An excellent history of the BPA is presented in Shapiro (1989).

⁵When the BPA was established, less than 10 percent of the farms in the United States had electricity; by 1960, in contrast, over 97 percent had electricity, and today over 99 percent do.

⁶By 1932, for example, the eight largest electric holding companies controlled 73 percent of the privately-owned electric business. Holding companies which were severely condemned in a report by the Federal Trade Commission, eventually collapsed because of federal legislation prohibiting or discouraging anticompetitive activities. In 1935 Congress enacted the Public Utility Act, which was comprised of two parts. The one was the Public Utility Holding Company Act, which helped to put into practice present-day public utility regulation in the United States; the other was the Federal Power Act, which gave the then Federal Power Commission authority over the pricing of wholesale power and transmission transactions.

It should be noted, however, that most experts during the 1920s and 1930s considered large integrated electric power systems to be the most economical form of utility structures. Many of the mergers during the 1920s involved the integration of several small operating companies into a few large holding companies. To say, therefore, that holding companies were detrimental to electricity consumers and society at large requires more than presuming they were because of abuses. Holding companies may have represented an economically sensible corporate structure at a time when the demand for electricity grew rapidly and capital requirements were significant (see Hausman and Neufeld 1991).

⁷See, for example, Joskow (1989), Kriz (1991), Ruff (1989), Sawhill (1989), and the Winter 1992 issue of *Regulation*.

the true cost of providing power, competitive power trades and other market arrangements have the potential over the long run to reduce the cost of power, increase service reliability, and take away the need for additional environmentally and economically costly power projects.

These fundamental changes suggest a clear need to reassess the BPA's original mandate to supply "affordable" electricity to preference consumers. As the BPA is presently organized, it lacks the incentive structure to operate in a commercially sound manner or in the long-term interests of electric power consumers and producers in the Pacific Northwest. Thus, while the BPA is directed by statute to price electricity economically and repay its debts to the federal government, it has failed to do so and cannot be expected to do so in the absence of market-based reforms.⁸

The BPA's failure to price electricity to reflect the cost of securing additional supplies has generated a number of adverse effects. BPA's long-standing practice of underpricing power encouraged overconsumption, discouraged conservation, and artificially stimulated the expansion of additional power capacity.⁹ As a consequence of having exhausted cheap hydropower by the 1960s, the BPA and other regional utilities had to turn to more financially and environmentally expensive thermal plants.

One of the more well-publicized outcomes of these actions was the \$2.25-billion default of the Washington Pacific Power Supply System (WPPSS) in 1983, the largest municipal default in history. Only two of the five planned (nonfederal) nuclear power projects were salvaged, and the BPA's acquisition of a large share of the generating capability made it responsible for a large portion of each project's costs, including debt service, whether or not they were completed (Cooper 1986 and Shapiro 1989).

⁸Under the Department of Energy Organization Act of 1977, the BPA and the other four power marketing administrations were officially brought under the control of the Department of Energy. The Federal Energy Regulatory Commission (FERC) became responsible for reviewing rates for all power sold within the region, and for transmission services; FERC, however, has limited authority. A 1987 ruling by FERC, for example, argued that the rates charged by the BPA to California utilities are based, *de facto*, on the policies and regulations of the U.S. Department of Energy. Consequently, FERC reasoned that it has no authority to review the rates BPA charged to California utilities ("Proposed BPA Intertie Policy Drawing Criticism," 1987).

⁹The Bush Administration's *National Energy Strategy*, published by the Department of Energy in 1991, similarly argues that the current pricing practices of federal power marketing administrations "discourage energy conservation and efficiency by underpricing electricity. This, in turn, increases demand for electricity and for more federal hydroelectric dams" (p. 38).

As of September 1991, the future principal and interest payments for all "nonfederal" projects of the FCRPS came to \$13.7 billion and may be considerably higher pending the resolution of ongoing litigation over unallocated project cost. This amount is exclusive of the over \$9-billion cumulative repayable investment the FCRPS owes the U.S. government.¹⁰ The appropriated debt dates back to the system's inception and was originally expected to be repaid to the U.S. Treasury within 50 years.

Historically, the interest rates on the FCRPS' debt have averaged around 50 percent below commercial rates. Randomly selecting seven projects completed by the FCRPS since 1960, the U.S. Office of Management and Budget calculated that the average interest rate for the projects was less than 40 percent of the average U.S. Treasury long-term interest rate at the time the projects went into service.¹¹

Because of this highly subsidized financing, the BPA's power rates do not reflect the full costs incurred in making the power available, and therefore electric power consumers and producers receive distorted signals for making consumption and production decisions. As of 1989, the FCRPS had paid only 14.7 percent of its federal debt, which from 1987 to 1989 grew by \$302 million.¹²

In addition to the considerable cost to the general taxpayer, the cost of distorted price signals to power consumers, the cost of discriminatory Intertie policies, and the cost imposed on the environment from construction of unnecessary projects, the BPA's noncommercial practices also ensure that it has an unfair competitive advantage over privately owned utilities. For example, as a federal agency, the BPA's income and properties have always been tax exempt. Thus, as it presently functions, the BPA prevents the Pacific Northwest from realizing the real benefits emerging under a competitive U.S. electric industry.

This paper assesses the social costs generated under the BPA's present structure and shows why they can be expected to grow in the future as the U.S. electric industry moves rapidly toward competition and de facto deregulation of wholesale power transactions. Three proposals for reform are offered that, by introducing competitive forces and shifting BPA's resources away from the federal

¹⁰U.S. Department of Energy (1992).

¹¹U.S. Office of Management and Budget (1991). The mean interest rate for the seven BPA projects was 2.8 percent, while the mean U.S. Treasury long-term interest rate was 7.5 percent. OMB calculated that the "top" 228 federal power marketing administrations' projects will cost the U.S. Treasury \$300 million annually or \$13 billion in interest subsidies over the duration of the outstanding loans (p. 4).

¹²U.S. Department of Energy (1991a).

government and toward local control, should benefit not only the BPA's wholesale preference consumers but other power consumers in the Pacific Northwest as well.¹³

The Economic Costs of the BPA

Overview of the Evidence

The best measure of the magnitude of social costs generated under the BPA's public ownership structure is found by examining BPA's pricing policies and the scope of the federal subsidies the agency has received. Although BPA's originating legislation stipulated that rate schedules shall be established to recover all costs including amortization of the capital investment over a "reasonable" number of years, an examination of its pricing and debt repayment practices indicates that the agency never took this mandate seriously.

Yvonne Levy (1980) calculated that if the BPA had to pay market-based interest rates and taxes that privately owned utilities are obligated to pay, its rates per annum would have been 82 percent higher during the period of 1947 to 1979, 112 percent higher during the period of 1965 to 1979, and 134 percent higher during the period of 1971 to 1979.¹⁴ David Shapiro (1989) calculated that normal, straight-line amortization of BPA's debt alone would have increased the cost of operation (or conversely, decreased U.S. taxpayer subsidies) by \$1.05 billion for the period of 1973 to 1986.

Subsidies to the BPA partly stem from its use of arbitrary debt repayment schedules (it does not have fixed payback periods), and its ability to extract funds from the federal government at below market interest rates. These subsidies run into hundreds of millions of dollars per year. Shapiro (1989) estimated that the total subsidy to the BPA for the period of 1973 to 1983 was about \$4 billion. Timothy Roth (1986) estimated that, as of the end of fiscal year 1972, the BPA had repaid just 19 percent of the total federal hydropower investments. According to Milton Copulos (1986), by 1986 the BPA had repaid only 8 percent of the federal investments it made from 1937 to 1986. Like the other federal power administrations, Bonne-

¹³Recent proposals to deal with the BPA problem focus on repaying BPA's debts to the federal government. Senator Mark Hatfield proposed for consideration the use of private financing to repay BPA's current debts with the U.S. Treasury. Similarly, O'Connor and Olson (1990) propose local ownership of the BPA's electric power system as a political solution to the debate over repayment reform and privatization.

¹⁴Levy also estimated that for 1980 the BPA sold power at an average price that was only 13 percent of its marginal cost.

ville was, by law, supposed to repay its investments within 50 years of start-up for projects financed with federal funds.

In 1985, then OMB Director David Stockman testified before a congressional committee that BPA would have to raise electric rates by 20 percent in 1986 with debt repayment reform.¹⁵ Andrew Kleit and Richard Stroup (1987) found that for the \$6.5 billion in loans outstanding in 1986, the average interest rate was only 3.5 percent.

Additional Costs of BPA's Noncommercial Operations

Although the above estimates of subsidies and required rate increases demonstrate that the social costs generated under BPA's present structure are significant, they are likely to be conservative since they do not take into account the wider consequences these uneconomical practices have had on BPA's consumers, competitors, and the environment.

From the perspective of society at large, subsidies tend to prevent the most efficient electricity suppliers from selling in wholesale or retail markets. Privately owned utility operators have expressed concern that a move toward a more competitive environment will place them at a competitive disadvantage if publicly owned and rural electric cooperatives continue to receive large subsidies from the federal government.

Ironically, the BPA's below-cost pricing policy has led to major problems for electricity consumers in the Pacific Northwest. By artificially stimulating demand for power, underpricing forced the BPA to look for new sources of electricity supply in larger amounts and earlier than it would have otherwise. Through the 1970s, the BPA and a consortium of publicly owned utilities formed the WPPSS to build five nuclear power plants. To get around the legislative prohibition against owning power plants, the BPA bought rights, through "net-billing agreements" with other utilities, to power from three of the plants.¹⁶

Because of these "net-billing agreements," the BPA acquired most of the generating capacity of the five nonfederal nuclear projects. These contracts made the BPA and its customers financially responsible for the projects' entire costs. The burden was felt particularly

¹⁵U.S. Office of Management and Budget (1985).

¹⁶Net billing allowed the BPA to own the long-term contracts for all of the power from nuclear units 1 and 2, and 70 percent of all power from nuclear unit 3. The BPA applied a credit to the prices charged WPPSS utilities to help pay the interest and the principal on bonds the utilities issued to fund the nuclear power plants. Funding for the credit came from other BPA customers in the form of higher power rates.

by aluminum smelters, whose electricity rates rose about 800 percent over the period of 1979 to 1984 (Spies 1990).

In 1983, after the incurrence of significant cost overruns, the WPPSS defaulted on \$2.5-billion worth of bonds, and two of three plants were mothballed. According to the BPA's annual financial report for fiscal year 1991, the future principal and interest payments required for nonfederal projects total \$13.7 billion, of which \$6.8 billion represents interest.¹⁷ Ongoing litigation over two of the plants could result in the BPA assuming an additional \$1 billion in costs (Spies, p. 33).

In retrospect, much of the blame for the WPPSS financial disaster can be traced to the BPA's below-cost pricing, its inaccurate electricity demand forecasts, (see e.g., Wenders 1986), and the fact that, as a government entity with access to taxpayer funds, it did not have to bear the burden of its actions. Moreover, because the BPA through its "net-billing agreements" shouldered nearly all the risks associated with the projects, the other participating utilities did not have a strong incentive to closely monitor the construction of the projects.¹⁸

The BPA-subsidized prices, in addition, also encouraged consumers to underinvest in energy conservation. Artificially low electricity prices naturally lead to under-investment in energy conservation. When BPA's prices started to rise sharply in the 1980s, residents and businesses in the Pacific Northwest suffered large losses because of their high dependency on low-priced electricity. The "high dependency" reflects the low electricity rates in BPA's service area relative to the rest of the country. For example, residential electricity rates in Washington, Idaho, and Oregon are the lowest in the country; residential consumers in the three states combined pay about 55 percent of the average U.S. price for residential electricity.¹⁹

Finally, because underpricing of electricity encourages consumption, increasing the need to expand power capacity, it unnecessarily imposes cost on the environment. The BPA's extensive system of hydropower dams on the Columbia River has contributed to the significant reduction in salmon runs by damaging the fish as they

¹⁷U.S. Department of Energy (1992, p. 31).

¹⁸For example, the net-billing agreement forced BPA customers as well as U.S. taxpayers to bear the risks of cost overruns, noncompletion, and other problems encountered by the Washington Pacific Power Supply System (WPPSS) in constructing the nuclear power plants. The BPA's obligations to the WPPSS were on a "take or pay" basis, where credit to WPPSS utilities would accrue regardless of whether their nuclear plants ever generated power. The BPA even lacked the authority to manage the construction of the WPPSS plants.

¹⁹U.S. Department of Energy (1989).

move out to the ocean and by impeding returning adult spawners (Butcher, Wandschneider, and Whittlesey 1986). By the early 1980s the number of salmon returning to the Columbia River Basin to spawn each year had dropped from a previous peak of 16 million to about 2.5 million.²⁰

The Emergence of a Competitive Electric Industry

The electric industry is becoming more competitive and pressures are mounting for deregulating specific components of the industry. Just a few years ago the electric industry was stridently opposed to competition and reform of traditional cost-of-service regulation. Ongoing technological, regulatory, and competitive changes, however, indicate that the industry is undergoing fundamental and irreversible restructuring.

The positions of many privately owned utilities have shifted toward acceptance of competitive forces over state regulation to determine their future financial fate. For example, many electric utilities now favor purchasing generating capacity from outside parties over building their own generating facilities and being subjected to unpredictable and inflexible regulation (Joskow 1989). In large measure this openness to markets is a direct response of the industry's experience with construction delays, cost overruns, canceled projects, and prolonged rate hearings that occurred under traditional public utility regulation.

Today, five fundamental changes are occurring in the electric industry that are restructuring it along competitive lines. First, it is expected that as much as 50 percent of new generating capacity through the year 2000 will come from nonutility suppliers. By the end of the 1980s, almost 30,000 MW of nonutility generating capacity was in place in the United States, and 40,000 MW was being developed.²¹ Several utilities have successfully conducted competitive power procurement programs to meet their future needs. Additionally, by making small generating facilities more economically attractive, technology has made the long-held presumption of economies of scale in generation invalid.²² Consequently, less capital-intensive,

²⁰"Protection for Fish to Cut Bonneville Power Output, May Boost Use of TPPs." *Electric Power Alert*, 22 January 1992, pp. 4, 5.

²¹U.S. Congress, Office of Technology Assessment (1989).

²²Until recently, technological changes favored large generating units. The trend in the industry toward smaller-scale and less capital-intensive generating facilities (for example, modular and staged combustion turbine, combined cycle) reflects a rational response to prevailing economic and political realities.

small generators with shorter lead times currently are able to compete with utilities having large generating facilities in regional markets where both have equal access to potential buyers.

Second, in response to competitive pressures, the industry is becoming less vertically integrated and more focused on the primary segments of the electric power generation, transmission, and distribution system. Horizontal integration through mergers and joint ventures has become more common as utilities search for ways to minimize production and transmission costs in the face of increasing competition (O'Connor, et al. 1988).

Third, new and existing industry participants are demanding more access rights to transmission systems that are owned and controlled by privately owned utilities (Houston 1991 and Costello 1988). With greater interconnection of utilities, consumers have more service choices. This trend indicates that increased competition is tied to transmission access. Some utilities are offering, or are proposing to offer, competitors easier access to their transmission system in return for more flexibility in the pricing of wholesale power.²³ As the number of sellers and buyers in wholesale power markets increases, political pressures over access to transmission lines should continue to intensify.

Fourth, pricing in the industry will be determined more by market conditions and less by traditional cost-of-service regulation.²⁴ Electric services will increasingly become unbundled, with consumers being offered a greater mix of prices and services. In a competitive environment, prices will more closely reflect the marginal cost of generating and delivering additional power.

Lastly, further de facto deregulation of wholesale power markets and other workably competitive markets is likely to take place over the next several years. As the number of buyers and sellers increases and the benefits from trading increase, the regulatory system will come under considerable pressure to change and accommodate the growing interests who will gain under a more open and competitive environment. Support for a more competitive electric industry also

²³For example, Public Service of Indiana received approval from the FERC to price 450 megawatts of wholesale electricity at market prices in return for granting both utility and independent suppliers greater access to its transmission system (Federal Energy Regulatory Commission Docket No. ER-89-692).

²⁴In a competitive environment, utilities and other suppliers would be pressured to apply market pricing to both existing and potential new customers. For a utility threatened with "uneconomical bypass," market pricing becomes necessary in preventing customers from switching to higher cost suppliers or technologies. Market pricing increases benefits for society at large and a utility's profits, as well as reduces the need to raise rates to so-called "captive" customers.

is reflected in the Energy Policy Act of 1992. The act will assist in expanding access to electricity transmission for wholesale participants and in lifting key entry restrictions into wholesale power generation.²⁵

Increasingly, FERC has allowed transactors more freedom to sign contracts rather than being subjected to traditional rate-of-return regulation (Acton and Besen 1985, and Joskow 1989). As reflected in some of its major rulemakings issued in 1988, the FERC position has shifted to favoring competition over traditional regulation in determining the price of electricity.²⁶ FERC's liberalization of pricing rules for wholesale services and its policy shift toward lifting regulatory restrictions on suppliers of coordination and transmission services reflect its recognition of the important role competition has to play.

Even state regulators, by endorsing competitive bidding of new generation capacity and market-based rates for consumers who threaten to bypass the local utility, have become receptive to the increasingly competitive conditions in the electric industry. Since 1986 several state regulators have allowed electric utilities to offer special rates to customers who threaten to install cogeneration facilities (Costello 1989). In addition, as of March 1990, competitive power procurement programs operated in twenty-six states (Rose, et al. 1991).

The electric industry's increasingly competitive and demand-responsive character could offer significant benefits in the form of greater access to lower-cost electric power to the Pacific Northwest. The BPA's current structure and its federal agency status, however, make it incompatible with encouraging this beneficial trend.

Proposals for Reform

The fundamental changes taking place in the U.S. electric industry and BPA's growing demand for additional power make it particularly important that the BPA begin changing its structure and practices. Short of transferring all of BPA's assets and responsibilities to the private sector, a number of other reforms, designed to introduce the discipline of competitive markets, could improve the BPA's perfor-

²⁵The Energy Policy Act (P.L. 102-486), signed into law on October 24, 1992, is the first major federal energy legislation enacted in about fifteen years.

²⁶FERC issued three Notices of Proposed Rulemakings (NOPRs) on March 16, 1988 (Federal Energy Regulatory Commission Docket Nos. RM-88-000-4, 5, and 6), addressing independent power producers (IPPs), avoided cost rates for purchased power, and all-source bidding for purchased power, respectively. Although these rules were never enacted, FERC has carried out some of their provisions on a case-by-case basis.

mance in providing reliable, efficiently priced power. To be politically viable, any reform proposal must not only increase the efficiency of the overall system, but also compensate those who benefit from the current arrangement. Essentially, this requires that preference customers suffer no or minimal economic losses. The proposals presented below attempt to achieve this outcome.

Reselling Preference Electricity

Allowing the reselling of electricity by preference customers would represent a seemingly simple but important reform that could eliminate one major source of inefficiency of the BPA system. This positive outcome stems from the Coase-theorem idea that, when transaction costs are minimal, efficiency is achieved whenever participants bear the full social cost of their actions. To put it differently, efficiency does not depend on the initial allocation of property rights in a world without transaction costs (Coase 1960).

Currently, the BPA prohibits private persons or agencies, with the exception of privately owned utilities, from reselling electricity to privately owned utilities. The economic losses from prohibiting power resale reflect the difference between the market value of "preference" electricity and the value placed on it by preference customers themselves, who are the original recipients of the electricity. Reselling would therefore achieve roughly the same efficiency gains as would auctioning off all the available hydropower in a world where all preference rights are eliminated.²⁷

Under one proposal that preserves existing preference rights, the original preference recipient would have the option of either consuming the low-priced electricity marketed by the BPA it is currently entitled to or reselling any portion of it in the marketplace. A good analogy is the current trend in water reallocation in the West, where farmers with preferential rights to cheap water are being given the rights to resell the water they conserve at a market-based price up to their traditional allocations (Anderson 1983). Preference customers would have an incentive to resell whenever the value they place on electricity lies below the value placed by market bidders. The cost of energy conservation illustrates one possible value that preference customers, at the margin, may place on the electricity they consume. A privately owned utility may assign a higher value to a preference customer's electricity if the alternative is to build an expensive new power plant.

²⁷The major difference in outcome between the two arrangements (assuming minimal transaction costs) lies with the distribution of gains; unlike reselling, auctioning off all hydropower at market prices would directly benefit U.S. taxpayers.

Reselling would also promote energy conservation and, at the same time, protect the environment from unnecessary damage. By reselling in a competitive market, the preference customer would more efficiently consume electricity since the opportunity cost of electricity consumption to the customer would reflect the true market value of electricity rather than the subsidized, below-market price the customer currently pays the BPA. By facing the real market price for electricity that would more accurately reflect the value of the resource, preference consumers would have better information and incentive with which to make efficient consumption decisions. Moreover, by encouraging conservation and trading of power, reselling would diminish the reliance on financially and environmentally costly new generating capacity.

As a major outcome, reselling electricity would benefit the original recipients, namely preference customers and their retail customers.²⁸ If the original recipients decide to resell the electricity, they would be better off economically than if they lacked the right to resell. But just as importantly, by facilitating exchange, another party would also benefit; namely, a buyer who is able to purchase a source of valuable electricity whether because it is lower-priced, more conveniently available, or more attractive in some other way than alternate sources of supply.

One method of facilitating the exchange of preference power would involve the federal government soliciting bids from all interested purchasers specifying their willingness to pay for different quantities of power. The auction would include all of the hydropower available to both preference and nonpreference customers. Taking into account individual bids, the buyer's willingness to pay for available power can be measured. The market price would be determined by the interaction of bidders' willingness to pay with the total available power offered by the federal government. Bidders without preference rights would pay the market price. Preference customers also would pay the market price for the electricity they wish to consume, but would receive a credit equal to the market price times their original rights to power.

To illustrate the above mechanism, assume that a preference customer is willing to buy 500,000 kilowatthours of electricity at the market (bid) price of five cents per kilowatthour; assume also that the customer has an original preference right to one million kilowatt-

²⁸Preference customers typically are power distributors owned by the retail customers they serve. The average customer would therefore stand to benefit from the economic gains of resale.

hours. Consequently, the customer would pay \$25,000 (five cents times 500,000 kilowatthours) for electricity it wishes to consume and, concurrently, would receive a credit of \$50,000 (five cents times one million kilowatthours) for its original rights. The preference customer, on net, receives \$25,000 for allocating 500,000 kilowatthours of its original rights to the market. In other words, the preference customer receives the market value for electricity that it makes available to others.

Customers would gain from reselling whenever they could be compensated for the unused electricity, for example, by "buying" energy conservation at a cost of less than five cent per kilowatthour. Under the current regime, the preference customer would tend to only conserve when the cost of conservation is less than the subsidized power prices charged by the BPA. If the preference price equals three cents per kilowatthour, for example, and the cost of conservation equals four cents per kilowatthour, the customer would rather consume than conserve since conserving is more expensive than buying electricity.

If instead the market price was five cents per kilowatthour, the preference customer would be better off by reselling electricity and purchasing conservation since the price received for selling electricity is greater than the cost of offsetting the sale by conservation. Under this arrangement, preference customers would tend to have an incentive to resell any time the market places a higher value on electricity than what they do individually. Preference customers would reduce their electricity consumption up to the point where the cost of conservation is equal to the market value of their allocated electricity. Reselling, then, would give preference customers a stronger incentive to conserve electricity. Currently, because of the prohibition on reselling, preference customers overconsume electricity at artificially low prices, and nonpreferential consumers underconsume at artificially high prices.

In sum, the auctioning off of excess preference rights has two major benefits. First, preference rights would be transferable to the highest-valued uses. In the above example, if the preference customer values its allocated electricity at less than five cents per kilowatthours—the market price—it would have an incentive to sell to a buyer who values electricity higher. Second, by allowing any party to bid on electricity that preference customers are willing to offer at the market price, auctioning off of excess preference rights would stimulate competitive conditions, and the price of electricity should therefore more closely reflect the cost of providing additional power.²⁹

²⁹The economic benefits and political palatability of reselling preference power in the Pacific Northwest also are discussed in Kleit and Stroup (1987).

Sale of BPA's Transmission System

Sale of the BPA's transmission line to interested buyers would promote competition in the Pacific Northwest electric industry and increase the efficiency gains from reselling. The recent debate over competition in the U.S. electric industry has focused on the rights of independent power producers and customers to gain fair access to transmission networks.

Competitive conditions require that privately owned utilities, independent power producers, publicly owned distributors, rural electric cooperatives, and other participants in the electric industry have access to the transmission network. A 1989 study by the Office of Technology Assessment concluded that, in competitive wholesale markets, technical conditions should not hamper the increased demands that would be placed on transmission systems.³⁰

In the absence of nondiscriminatory access, some interested entrants may face difficulties acquiring financial capital or selling their electric services in a spot market, or under long-term contracts, to nonlocal buyers.³¹ In addition, distributors and other purchasers would be deprived of the lowest cost, available electricity in a regional market.³²

The major policy issue revolves around the question of what is the most efficient institutional arrangement for giving "nonBPA" groups access to the transmission network: How much, and at what level of, control of the current and future BPA transmission network do the various suppliers and consumers located in the Pacific Northwest need to foster competitive conditions?

The major argument for joint ownership of BPA's transmission network centers on the need to establish well-defined rights that would give users incentives both to utilize the network efficiently and to invest in new capacity when warranted by market conditions. By possessing the rights of access, exclusion, and transferability, users would realize maximum benefits from investing funds in the

³⁰Congress of the United States, Office of Technology Assessment (1989).

³¹Under many circumstances, the spot market can shift risks efficiently. In a spot market the risks of planning and operation fall on investors rather than consumers. A spot market can also promote efficiency by setting prices that correspond to actual market conditions.

³²"Shopping around" induced by fair transmission access should improve both allocative and productive efficiency: sellers would be pressured by competitive forces to offer prices corresponding to their marginal costs, or else lose customers to other suppliers; it is also more likely that electricity consumers would be served by the lowest-cost suppliers since consumers would not be constrained to purchase their electricity from local suppliers (see Costello 1988).

network as a response to new demands brought on by a growing market.

The benefit to society of joint ownership stems from the likelihood that the transmission assets currently owned by the federal government will have a higher value in the marketplace than what they now have. This is because joint ownership would create new incentives stimulating economical investments in new capacity and efficient use of existing capacity. Individual owners would have the right to lease or sell their portion of the transmission network to outside parties or other existing owners of the network.

One proposal would involve the BPA auctioning off capacity that is not needed to serve preference customers and other customers with which the BPA has existing service obligations. The funds from the sale could be used to repay BPA's outstanding debts to the U.S. Treasury. For example, the value of BPA's transmission system is roughly \$10 billion; this equals about 65 percent of BPA's total outstanding debt to the U.S. Treasury and other bondholders.³³ Alternatively, the BPA could offer to sell parts of existing capacity to preference customers at a discounted price. In either case the current BPA transmission network would evolve into a joint privately owned/publicly owned network.

The newly formed regional transmission company would devise rules giving managers wide discretion to operate the system in a way necessary for maintaining technical integrity. Rules also would require that new users be allowed to purchase and to receive the same rights as current users. Such rules would prevent existing owners from foreclosing the entry of new generators into the regional power market. The new users would have to abide by the operational rules established by the transmission company and meet their financial obligations for funding new capacity and maintaining current capacity.³⁴ Without these requirements the value of current and future assets would likely fall, thereby discouraging current owners from making additional investments.

The owners collectively would form pricing rules falling under the scrutiny of antitrust laws. It is expected that, under newly developed competitive conditions, owners-users would refrain from price-fixing practices since it would not be in their interests to establish arti-

³³See U.S. Department of Energy (1991a, p. 10) and (1992, p. 21).

³⁴Separation of ownership rights and operation rights currently exists for the National Grid Company, which is the newly formed British common-carrier, electric transmission utility. The company is owned by twelve regional electric distribution companies.

cially high prices that would lower each owner's share in the wholesale power market.

Joint ownership by regional users of the transmission network has distinct advantages over other proposals such as contract-common carriage and incentive-based pricing of transmission services.³⁵ Joint ownership would: (1) avoid the costs of monopoly power currently being exhibited by the BPA, reflected in the priority given to power generated and sold by the federal government; (2) transfer ownership-control rights among parties on the basis of economic value since access would be available to those who value it the most, as reflected in the prices they are willing to pay for transmission service or for ownership rights in the regional transmission company; (3) eliminate transaction costs that would otherwise be incurred in negotiating complex contracts with the BPA; (4) eliminate the need for complex access and pricing rules that would induce costly legal and judicial interference (antitrust enforcement and FERC oversight could ensure that the regional transmission company does not artificially restrict the entry of new owners); and (5) eliminate the ambiguity of ownership rights and governance under the present arrangement, and thus reduce legal and political costs in addition to encouraging efficient use of the transmission network.

The BPA currently controls about 80 percent of the Pacific Intertie extending from the Northwest to California.³⁶ The BPA faces little oversight by FERC and is exempt from antitrust laws and therefore is able to discriminate against privately owned utilities and other groups to favor its preference consumers. BPA's Intertie pricing and access policies over the years have met with bitter opposition by regional and nonregional electricity producers and consumers. California utilities and regulators for years had protested against BPA's Intertie access policy as both anticompetitive and discriminatory. Specifically, they have charged that the BPA sells power to the California market at discriminatory prices, reflecting a BPA policy of minimizing revenues that the BPA needs to collect from its preference customers.³⁷

³⁵For problems associated with contract-common carriage, see Houston (1991). "Better pricing" may be a remote possibility in a world where the BPA continues to have ownership and control rights of the transmission network. Political considerations and exploitation of BPA's market power make it highly unlikely that prices would be set at efficient levels.

³⁶The combined capacity of the Pacific Intertie equals 5,156 megawatts of which 4,056 megawatts fall under the BPA's control.

³⁷See "Proposed BPA Intertie Policy Drawing Criticism" (1987, pp. 36-37), and Tussing (1988, pp. 111-19).

The major economic constraint under the current single-ownership regime is that new competition in alternative transmission lines is impractical because individual generators are severely restricted from constructing their own transmission lines over different rights of way. In contrast, under joint ownership, competition would be directed to vying for ownership and control rights in existing transmission lines. Thus, even in the presence of scale economies, competitive conditions would still prevail by allowing groups to sell or lease rights in transmission capacity to others.³⁸

Finally, the joint owners would perform the same functions as the BPA currently does, including planning for and financing new transmission capacity, operating with sufficient capacity and system interconnections, and coordinating electricity flowing from different generators to their purchasers.³⁹ Participants in the electric power market would gain essential rights that they currently do not have, namely rights for market-based access and prices. In contrast to the BPA's present Intertie policy, joint ownership would promote competition in ownership control and provide strong incentives to coordinate power efficiently and to make economically sound capacity expansion decisions.

An End to Centralized Government Energy Planning

The Pacific Northwest Electric Power Supply and Conservation Act of 1980 contains a provision that can diminish the efficiency of the region's electric power industry.⁴⁰ The provision requires the

³⁸ Arguments for joint ownership of electric transmission lines also are presented in Houston (1991), Costello (1988), and Smith (1987).

Alger and Toman (1990) have proposed joint ownership of natural gas pipelines to promote competition in the natural gas markets. According to the authors, joint ownership would allow pipelines to "reap scale economies by competitive offering of transportation services by multiple shareholders to limit the exercise of market power" (p. 276).

³⁹ Joint ownership must reckon with three major technical conditions required for a smoothly functioning electric power system. First, the electricity flowing from individual generators to purchasers and from different generators, all using the same transmission network, would have to be coordinated. Second, the regional transmission company would need to take into account the interdependency between the three primary components (generation, transmission, distribution) of an electric power system. The workings of the three components are intertwined in a way that allows an electric power system to operate as a coordinated whole. Third, the flow of electricity on outside electric power systems should be held to a tolerable level or, if not possible, compensation should be paid to utilities located outside the intended transmission path.

⁴⁰ The act was a response to growing threats, starting in the 1970s, of rising energy prices and anticipated power shortages in the Pacific Northwest. These threats jeopardized the future availability of low-priced power to privately owned utilities and direct service industries located in the Pacific Northwest. Politically, the act tried to appease all electricity consumer groups by expanding the BPA's role in the Pacific Northwest

newly formed Northwest Power Planning Council to give priority to energy conservation as a new resource to meet future demand needs.

The act applies a 10 percent premium to energy conservation by discounting its cost for environmental advantages when placed side-by-side with traditional resource alternatives. The act also gives priority to renewable resources and generating resources applying waste heat or high fuel conversion efficiency over other resources (for example, gas-fired facilities).⁴¹

In accordance with the act, the BPA is required to take actions that are consistent with the approved plan of the Northwest Power Planning Council. The Council is composed of representatives from Oregon, Washington, Idaho, and Montana, and its primary job is to develop a twenty-year power plan for the Pacific Northwest. In every approved plan, the emphasis has been on promoting "cost-effective" energy conservation. The Council defines "cost effective" as a condition under which the cost of energy conservation is less than the cost of new supply resources.⁴²

A basic premise of the Council's planning is its perception that energy conservation should be subsidized at the cost of new supply resources (for example, base-load coal plants); the Council believes that energy-conservation actions subsidized up to the cost of new resources are "cost effective." To the contrary, however, subsidies based on the cost of new supply resources provide a consumer with excessive incentive to conserve electricity: the opportunity cost to the consumer of using more electricity includes both the price of electricity and the cost of new supply resources (that is, the subsidy offered to the consumer for using less electricity). The consumer's opportunity cost from using additional electricity therefore would exceed the marginal cost (see e.g., Joskow 1988 and Kahn 1992).

To illustrate this point, if the price of electricity equals three cents per kilowatthour and the cost of new supply resources (the marginal cost) equals five cents per kilowatthour, a consumer would "pay" eight cents for using an additional kilowatthour of electricity; that is, the eight cents represent the opportunity cost to the consumer from using more electricity. The consumer therefore pays a real price of eight cents for using electricity, while the real cost to society is only five cents. From society's perspective, the consumer would have an incentive to use too little electricity.

power market.

⁴¹See Northwest Power Planning Council (1989, p. 4).

⁴²See Northwest Power Planning Council (1989, p. 4).

In addition, such subsidies would harm nonrecipient consumers, since BPA's distributors receive less revenues and make subsidy payments equal to the cost of new supply resources. To make nonrecipient consumers no worse off, the maximum subsidy should not exceed the difference between the cost of new resources and the current price of electricity. Instead the BPA subsidies require nonrecipient consumers to help pay for conservation investments that would generally benefit only a small minority of a preference customer's retail buyers.

The Northwest Power Planning Council apparently believes that the pricing system is deficient at achieving the socially efficient amount of conservation because electricity consumers are either irrational or ill-informed, or both, when making decisions on how much energy conservation to purchase. But, evidence of market imperfections serious enough to justify subsidies for energy conservation has not emerged. Studies that allege the presence of market imperfections fail to account for factors such as consumers expecting to earn higher returns from other types of investments, the high uncertainty of actual electricity savings, and the transaction costs associated with purchasing conservation measures (see e.g., Newlong and Weitzel 1991).

Ironically, by failing to price electricity at its real cost, the BPA not only undermines the most effective mechanism for fostering conservation but, by seriously underpricing power, it actually encourages consumers to overuse electricity and underinvest in conservation.⁴³

The most effective way for BPA to promote energy conservation in the Pacific Northwest is to set the price of electricity at its market value. The right to resell preference power, the first reform offered in this paper, would achieve the same objective by informing preference users of the real value of the power they consume.

The Northwest Power Planning Council's experience indicates why government attempts at energy planning are often likely to be misguided and counterproductive. By suppressing the crucial role prices play in guiding consumption and investment decisions, subsidies, whether for below-cost power or for energy conservation, impede the ability of consumers to make cost-minimizing energy decisions. At the same time, by distorting market prices, subsidies

⁴³Substitutes for electricity in various uses also may be underconsumed. For example, many households in the Pacific Northwest may decide to switch from electricity to natural gas for water heating if BPA's prices were more comparable with economic costs.

impede efficiency-enhancing competition in the supply of electricity. By making power artificially cheap, subsidies in the Pacific Northwest have likely discouraged the development of independent power production, which may in fact be the most cost-effective source of new supply.

Beneficiaries

Taken together, the three reforms advanced would go a long way to benefit the nation, the Pacific Northwest, and various interest groups affected by the BPA's current practices. They offer environmentalists a way to curb the unnecessary expansion of new generating capacity, including hydropower facilities and fossil fuel plants, induced by politically based pricing practices. To conservationists the proposals promise to offer electricity consumers in the Pacific Northwest and adjacent regions more opportunities to make efficient conservation investments. For privately owned utilities the proposals open new markets for both the electricity they sell and the electricity they buy. U.S. taxpayers would benefit by allowing BPA to reduce its debt with the U.S. Treasury from the sale of transmission assets and by seeing that BPA's operations are put on a sounder financial basis. Independent power entrepreneurs would gain by having more opportunities to compete fairly with power marketed by the BPA. To public power entities and their customer-owners in the Pacific Northwest the proposals offer the choice of whether to continue consuming the power they are currently allocated or to resell a portion to interested buyers at a profit.

Most importantly, the reforms would benefit the nation as a whole by stimulating competition in the electric industry and by promoting efficient pricing and trading. Competition almost assuredly would benefit the long-term interests of both electricity consumers and the nation.

Finally, the proposals would eliminate wasteful political costs currently being expended by special interest groups in their quest for low-cost hydropower. It is questionable how much preference customers have benefited since the inception of the BPA. Because of the constant battles involving high litigation and lobbying costs to maintain preference status, in addition to the inefficiencies of the BPA, much of the initial benefits received by preference customers may have disappeared over time.

Conclusion and Policy Implications

The future of the Bonneville Power Authority should be assessed in light of the competitive and efficient market-oriented arrange-

ments that are increasingly emerging in the U.S. electric industry. Because as a government entity it does not bear the costs of its actions nor is capable of being positively influenced by competitive forces, the BPA's present structure is incompatible with this emerging trend.

The federal government can make three fundamental reforms which, by making the agency's practices more compatible with market incentives, would improve the BPA's performance and allow it to participate in the economically beneficial changes occurring in the rest of the electric industry.

First, the BPA's preference customers should be allowed to resell electricity to those who value it more highly. Second, the BPA's transmission system should be privatized gradually over time to reduce the agency's market power over the system and to facilitate access to the system. Third, the BPA should rely on efficient market pricing of electricity rather than special treatment for conservation for promoting the wise use of energy.

By placing greater reliance on market arrangements and less on federal control, these reforms have the merit of both increasing efficiency and accountability, whereby those groups most affected by BPA's policies and practices will be in a better position to influence its direction. By making the BPA more responsive to market incentives, these proposals will make the BPA's actions more compatible with the competitive changes sweeping the rest of the U.S. electric industry.

References

- Acton, Jan Paul, and Besen, Stanley M. *Regulation, Efficiency, and Competition in the Exchange of Electricity: First Year Results from the FERC Bulk Power Market Experiments*. Santa Monica, Calif.: Rand Corporation, October 1985.
- Alger, Dan, and Toman, Michael. "Market-Based Regulation of Natural Gas Pipelines." *Journal of Regulatory Economics* 2 (September 1990): 263-80.
- Anderson, Terry L. *Water Crisis: Ending the Policy Drought*. Baltimore, Johns Hopkins University Press, 1983.
- Butcher, Walter; Wandschneider, Philip; and Whittlesey, Norman. "Competition between Irrigation and Hydropower in the Pacific Northwest." In *Scarce Water and Institutional Change*, pp. 25-66. Edited by Kenneth D. Frederick. Washington, D.C.: Resources for the Future, 1986.
- Coase, Ronald. "The Problem of Social Cost." *Journal of Law and Economics* 3 (October 1960): 1-44.
- Cooper, Peter D. "The Bonneville Power Administration: The Worst Mess by a Dam Site." *Cato Policy Analysis* no. 66. Washington, D.C.: Cato Institute, 6 February 1986.
- Costello, Kenneth W. "Incentive Rates or Market Rates: A Rose by Any Other Name?" *The Electricity Journal* 2 (August/September 1989): 42-51.

- Costello, Kenneth W. "The Struggle over Electricity Transmission Access." *Cato Journal* 8 (Spring/Summer 1988): 107-24.
- Copulos, Milton R. "Cutting the Deficit by Selling Federal Power Marketing Administrations." Backgrounder no. 485. Washington, D.C.: The Heritage Foundation, 13 February 1986.
- Hausman, William J., and Neufeld, John L. "Public Policy and the Structure of the Electric Power Industry: Lessons from the Past." Paper presented at the 23rd Annual Conference of the Institute of Public Utilities, Michigan State University, Williamsburg, Va., 10 December 1991.
- Houston, Douglas A. "Toward Resolving the Access Issue: User Ownership of Electric Transmission Grids." Policy Insight no. 129. Santa Monica, Calif.: Reason Foundation, August 1991.
- Joskow, Paul L. "Regulatory Failure, Regulatory Reform, and Structural Changes in the Electrical Power Industry." In *Brookings Papers on Economic Activity: Microeconomics 1989*, pp. 125-208. Edited by Martin Neil Baily and Clifford Winston. Washington, D.C.: The Brookings Institution, 1989.
- Joskow, Paul L. *Testimony before the Subcommittee on Energy and Power*. House. Committee on Energy and Commerce, Washington, D.C., 31 March 1988.
- Kahn, Alfred E. "Least Cost Planning Generally and DSM in Particular." *Resources and Energy* 14 (1,2) (April 1992): 177-85.
- Kleit, Andrew N., and Stroup, Richard L. "Blackout at Bonneville Power." *Regulation* 11 (2) (1987): 30-36.
- Kriz, Margaret E. "Power Struggle." *National Journal*, 6 April 1991, pp. 789-93.
- Levy, Yvonne. "Pricing Federal Power in the Pacific Northwest: An Efficiency Approach." *Federal Reserve Bank of San Francisco Economic Review* (Winter 1980): 40-63.
- Newlong, T. Scott, and Weitzel, David. "Do Market Imperfections Justify Utility Conservation Programs? A Review of the Evidence." *The Electricity Journal* 4 (4) (May 1991): 40-53.
- Northwest Power Planning Council. *Supplement to the 1986 Northwest Conservation and Electric Power Plan, Volume 1*. Portland, Oreg.: Northwest Power Planning Council, 1989.
- O'Connor, Philip R.; Levin, Stanford L.; Cahill, James J.; and Keenan, Gerald M. "The Transition to Competition in the Electric Utility Industry." *Journal of Energy Law and Policy* 8 (2) (1988): 222-54.
- O'Connor, Philip R., and Olson, Wayne. "Localizing Federal Power Assets: A Path for the Pacific Northwest." *Public Utilities Fortnightly*, 1 February 1990, pp. 7-13.
- "Proposed BPA Intertie Policy Drawing Criticism." *Public Utilities Fortnightly*, 15 June 1987, p. 37.
- "Protection for Fish to Cut Bonneville Power Output, May Boost Use of IPPs." *Electric Power Alert*, 22 January 1992, pp. 4, 5.
- Regulation* 15 (1) (Winter 1992). Special issue on *Electric Power Deregulation*.
- Rose, Kenneth; Burns, Robert E.; and Eifert, Mark. *Implementing a Competitive Bidding Power for Electric Power Supply*. Columbus, Ohio: The National Regulatory Research Institute, 1991.

BONNEVILLE POWER ADMINISTRATION

- Roth, Timothy P. "Power Marketing Administrations: The Case for Privatization." Washington, D.C.: Citizens for a Sound Economy, 31 January 1986.
- Ruff, Larry E. "Electricity Restructuring in Two Nations: Different Paths to a Competitive Future." *Public Utilities Fortnightly*, 22 June 1989, pp. 13–19.
- Sawhill, John C. "Planning to Win in the New Utility Environment." *Public Utilities Fortnightly*, 12 October 1989, pp. 19–33.
- Shapiro, David L. *Generating Failure: Public Power Policy in the Northwest*. Cato Institute Book. Lanham, Md.: University Press of America, 1989.
- Smith, Vernon L. "Current of Competition in Electricity Markets." *Regulation* 11 (2) (1987): 23–29.
- Spies, Paul. "Variable Electricity Prices for Aluminum Smelting in the Northwestern USA." *Energy Policy* 18 (2) (March 1990): 162–69.
- Tussing, Arlon R. "Electricity and Gas: The U.S. West." *Energy Journal* 9 (4) (October 1988): 111–19.
- U.S. Congress, Office of Technology Assessment. *Electric Power Wheeling and Dealing*. Washington, D.C.: Office of Technology Assessment, May 1989.
- U.S. Department of Energy. *Form EIA-213*. Washington, D.C.: U.S. Department of Energy, 1989.
- U.S. Department of Energy. *Bonneville Power Administration: Financial Summary 1990*. Portland, Ore.: Bonneville Power Administration, 1991a.
- U.S. Department of Energy. *National Energy Strategy: Powerful Ideas for America*. First Edition 1991/1992. Washington D.C.: U.S. Department of Energy, 1991b.
- U.S. Department of Energy. *Bonneville Power Administration: Financial Summary 1991*. Portland, Ore.: Bonneville Power Administration, 1992.
- U.S. Office of Management and Budget. "Fact Sheet on Reform of the Federal Power Marketing Administration's Debt Repayment Practices." Internal Memo. (April 1991): 3–4.
- U.S. Office of Management and Budget. *Testimony of the Honorable David A. Stockman, Director, Office of Management and Budget, before the Committee on Interior and Insular Affairs*. House. Subcommittee on General Oversight Northwest Power and Forest Management, Washington, D.C., 1985.
- Wenders, John T. "Efficiency, Subsidy, and Cross-Subsidy in Electric Utility Pricing." In *Electric Power: Deregulation and the Public Interest*, pp. 307–36. Edited by John Moorhouse. San Francisco: Pacific Research Institute for Public Policy, 1986.

MORALITY AND MONOPOLY: THE CONSTITUTIONAL POLITICAL ECONOMY OF RELIGIOUS RULES

Gary M. Anderson and Robert D. Tollison

Introduction

Moral philosophy and economic analysis have long been thought to be entirely separate and distinct undertakings. The modern theory of human capital has opened communication between these traditionally disparate areas. Moral character can be analyzed as a form of reputational human capital. An individual can expect to bear significant personal costs from a moral transgression, to the extent that his or her reputation as a trustworthy, morally responsible agent is jeopardized by the immoral act.

Unfortunately, moral rules are not entirely self-enforcing in this manner. Some moral rules tend to benefit society, even though the individuals following those rules may not themselves directly benefit. This dichotomy presents an intriguing problem.

Constitutional economics suggests that moral rules form a relevant portion of the infrastructure of economic society. Some moral rules constrain the behavior of individuals in economically efficient ways, benefiting society at large. Such rules function to minimize free riding problems by providing internal constraints on individual behavior. The available stock of moral capital is a vital prerequisite for an efficiently functioning market economy.

Throughout history, a major provider of this moral infrastructure has been organized religion. Adherence to moral regulations has for

Cato Journal, Vol. 12, No. 2 (Fall 1992). Copyright © Cato Institute. All rights reserved.

Gary M. Anderson is Professor of Economics at California State University, Northridge, and Robert D. Tollison is the Duncan Black Professor of Economics and Director of the Center for Public Choice at George Mason University.

The authors wish to thank Pamela Brown, William Brown, James Buchanan, Keith Evans, Richard Friedman, Robert Harding, Ivan Johnson, Hartmut Kliemt, Shirley Svorny, Viktor Vanberg, and Ben Yu for helpful comments. The usual caveat applies.

most individuals been related to their beliefs in the demands of theology. Stated simply, the belief in heaven and hell (or their equivalents in non-Christian religions) has been an important source of moral behavior.

Religious doctrine can be modeled as a kind of "economic good," provided to its "consumers" across a "market." By doctrine we mean both the system of theology per se, and the moral guidelines promulgated in connection with those theological beliefs. When confronting situations involving other, more ordinary, sorts of economic goods, we routinely assume that the structure of the market will have significant effects on the price, quantity, and quality of service provided. This paper examines the relation between market structure and morality, considered as the "output" of organized religion. We will argue that there is an unusual relationship between market structure and social welfare in this particular case. We normally expect monopoly provision of a good or a service to reduce social welfare. But monopoly in the market for organized religion may actually increase net social welfare.

Religion and the Provision of Moral Order

The modern economics of crime and law enforcement emphasize the measurable, and hence observable, costs and benefits associated with nonvoluntary transactions. The individual criminal engages in unlawful behavior only when he perceives the expected marginal benefit to exceed the anticipated marginal cost. This formulation, however, is seriously incomplete. After all, the actual risk associated with most forms of criminal behavior is probably quite small relative to the marginal benefits the criminal might expect to receive. For example, rates of apprehension, prosecution, and conviction of criminals are fairly low. Yet most people, most of the time, abstain from criminal acts; most transactions between individuals are voluntary.

Adam Smith noticed this apparent paradox, and concluded that part of the answer lay in the relationship between the income streams of individuals and their personal reputations for moral character. The profit opportunities available to a market participant are enhanced if he has a reputation for being a reliable, nonopportunistic potential transaction partner. This reputation is a valuable capital asset, and its increase in value rewards him for displaying honorable behavior. Therefore, the expansion of markets (and commercial interaction) tends to improve the moral character of individuals, i.e., those personality traits which tend to benefit society.

Adam Smith also argued that there was another critical constraint in addition to the market for reputational human capital. In his *Theory of Moral Sentiments* (1759), he described this internal constraint as the “impartial spectator.” Certain features of an individual’s own utility function might cause that person to prefer some forms of behavior and avoid others because of the predictable effects such acts might have on the welfare of other people.

Thus, Smith distinguished three distinct, but interrelated and mutually supporting, levels of constraints facing any individual’s behavior toward the rest of society: (1) external law enforcement, i.e., the machinery of justice in society; (2) external incentives for control of opportunistic behavior in the form of the market for reputational human capital; and (3) the internal structure of the individual’s utility function that performs self-monitoring against opportunistic individual behavior.

Douglass North (1981) has termed this last constraint “ideology”—using the term to refer not to a system of ideas but to systems of constraints internalized within individuals—and has argued that this internal self-monitoring function plays a vital role in motivating human economic behavior. David Gauthier (1986), David Levy (1982), and Viktor Vanberg (1988) all argue that this internalized self-monitoring process represents an important constraint on individual maximizing behavior that needs to be incorporated in economic models of individual action. Most writers refer to this internalized self-monitoring process as “morality.”

Moral rules function as internal constraints on individual behavior, which provide for self-monitoring and self-enforcement against socially inefficient forms of behavior. The totality of relevant moral rules accepted by all individuals represents the moral constitution of a society. The set of moral rules a person has internalized represents that individual’s personal moral constitution.

Vanberg (1988) argues that individual adherence to moral rules is motivated by two separate factors. Individuals have an incentive to protect the value of their reputational capital as trustworthy potential transactors (as noted by Adam Smith). But individuals also benefit from a personal moral constitution in that it allows them to minimize the costs of day-to-day decisions.

The moral choices of individuals also have repercussions for society as a whole. An efficient economic order will best operate in the context of an economically efficient set of moral rules. For example, assume that a moral rule allowing unrestricted theft (“it is moral to steal as long as you do not get caught”) maximized individual wealth in the short-run for successful thieves. Unfortunately, individual

moral choices that are efficient from the perspective of society may not be efficient from the standpoint of the choosing individual. To return to the above example, sometimes an individual might correctly perceive that the risk of being discovered as a thief is negligible, making a particular act of theft seem worth the cost, despite the fact that his thievery reduces the rate of return on productive investment and thereby reduces the net wealth of society as a whole. Moral choices made by individuals often imply this kind of public goods problem.

Some moral rules, held by different groups at various times in history, have undoubtedly hindered the operation of local markets. Competition between different societies that subsisted on different sets of moral rules—different moral constitutions—led to the systematic elimination, through natural selection, of moral rules that reduced economic efficiency and rewarded societies with different rules that contributed to efficiency. Therefore, over long periods of time, moral rules have tended to become more efficient because inefficient rules systematically failed to survive. Unfortunately, the “marketplace” of cultural evolution requires centuries, or even millennia, to clear.

Organized religions have provided most systems of moral rules accepted by individuals as the basis for systems of internal moral constraints. Moral rules are not necessarily religious in character, but systems of moral rules that individuals follow have most often been provided by religious organizations. Organized religions can be thought of as firms marketing systems of moral rules to individual consumers. Such “firms” enjoy economies of scale in the production and marketing of moral rules, and combine such provision with other advantages (e.g., the club benefits enjoyed by members of the same church).

Religion provides an additional level of enforcement to moral behavior: the supernatural “police.” Religions describe the systems of moral constraints they represent as expressing the will of the gods, or God, who threaten(s) to sanction malfeasance, either during the lifetime of the transgressor or in an afterlife (e.g., being consigned to hell rather than heaven). If omnipresent divine surveillance is a credible threat to potential sinners, fear of divine wrath may significantly raise the perceived cost of opportunistic behavior.

Throughout history, organized religions have collectively been a major provider of the moral constitution of society. The structure of the religious marketplace may have had significant effects on this process.

State-sponsored monopoly religion was roundly criticized by many classical liberals (e.g., Adam Smith and John Stuart Mill) as a violation of basic individual liberty, and also because the doctrines promulgated by such monopolies were thought to produce inferior moral constraints on individual behavior. Monopoly religions purportedly provide inferior services by tending to their flocks in a less diligent, more shirking manner. Social welfare suffers because the moral glue necessary for the efficient operation of the economic order is of lower quality. But competitive markets for religion and moral rules were thought to do a better job, much as competitive markets for any other good or service function to provide better quality at a lower price.

Consider, for example, Adam Smith's extensive discussion of the markets for religious services in *The Wealth of Nations* (see Anderson 1988). He argued that state-supported, monopoly religion produced inferior service for its consumers. Specifically, monopoly religion led to a relaxation in the effective moral standards promoted by clerics. This resulted from shirking on the part of clerics, who were left unconstrained by fear of competitive entry into the market for their services. According to Smith, monopoly in the market for religion tended to lead to a reduction in service quality (Smith [1776] 1976, p. 311).¹

Smith neglected to consider an important aspect of the behavior of a religious monopolist, and may therefore have reversed the sign on the net effect of monopoly religion on the moral constitution of society. Monopoly in the market for religious services may actually tend to *improve* the utility of moral rules, viewed from the perspective of society as a whole.

Morals and Competitive Markets

If morality is defined as a set of logically coherent, mutually consistent internal constraints on behavioral choices, then a large number of distinct ethical systems are possible. Alternative sets of moral rules are potentially available to rational individual actors. However, some conceivable moral rules contribute to social efficiency, while others may have the opposite effect.

Under the assumptions of the competitive model, we expect that consumers of religion will be provided with the precise mix of ser-

¹Smith ([1776] 1976, pp. 312–13) also suggested that a competitive market for religion might lead to excessive moral stringency from the standpoint of economic efficiency, as religious sects each attempted to convince consumers of their superior credibility by imposing more rigorous behavioral requirements on their adherents. He appears to have favored a basically free market for religion, with some minimal government regulation.

vices they consider optimal, subject to relevant cost constraints. The services of religion include the moral code associated with the religious belief. Religious firms will have a comparative advantage in the provision of sets of moral rules because the theological basis of such firms imparts credibility to the moral structure. A consumer has reason to rely on the quality of the given set of moral strictures because he believes that they have received divine sanction. This lowers the cost of marketing sets of moral rules to consumers and also lowers the cost of generating compliance. The same divine entity that guarantees quality provides for enforcement against moral transgressions (e.g., in the afterlife).

If religious doctrines achieve market success purely as a function of their ability to compete freely for adherents, the religions that survive will be those providing the most acceptable moral requirements. If we assume that all potential church members are members of the species *homo economicus*, then the best moral rules (from their individual perspectives) will be those that require them to do exactly what they would otherwise most prefer to do, other things equal. This does not necessarily imply that the list of rules will be an empty set; in fact, it might be consistent with immensely detailed moral strictures. Rather, it implies that however many nominal moral requirements there may be, they will collectively imply that the individual's behavior will be subject to no actual constraint over-and-above those constraints imposed by scarcity.

Moreover, in the long run, optimal sorting of individuals across available religions (each with its own peculiar moral strictures) by moral preferences will reinforce the nonbinding property of competitive moral rules. In other words, people will gravitate to religions that promote codes of morality producing the least impact on those particular individuals' actual behavior; people will tend to adopt systems of moral rules most consistent with their preexisting individual preference orderings.

The moral order—considered as the effect of, but separate from, the moral beliefs of individuals—is a pure public good. Rational wealth-maximizing individuals will act to protect the value of their reputations for being moral actors, but will tend to free ride and behave opportunistically in cases where they expect such malfeasance will remain undetected and unsanctioned. To a rational self-interested actor, behavioral (moral) restrictions are bads. Naturally, some individuals will be unaffected by some external restrictions. The costs of behavioral restrictions are a subjective phenomenon and may have no observable manifestation. But, in general, a rational individual will prefer fewer effective behavioral moral restrictions.

Competition in the market for religious services improves the efficiency of those markets in providing religious goods to consumers. These private goods include the insurance benefits (implicit and explicit) associated with church membership, the club gains, socializing, reputational advantages (Jane goes to church, so Jane is perceived to be more trustworthy by potential transactors as a result), and other effects. Such private goods will be available in larger quantities and at lower prices than would be the case if the religious market were monopolized.

But because consumers of different religions do not, other things equal, demand utility-reducing moral constraints to be imposed on themselves, moral order as a public good tends to be *underprovided* in a competitive market for religious services. Competing sects compete by improving service (and reducing price) to consumers—which includes producing and marketing more attractive (i.e., more liberal) moral restrictions. This process will tend to produce rule structures which are more congenial to those individual consumers.

This competition will not necessarily result in less stringent rules in any observable sense. Rather, a consumer of moral rules will prefer to join whatever religion offers a moral code of behavior that most closely resembles his or her underlying utility function. Thus, some religious sects or denominations may offer relatively complex and (superficially) strict sets of moral rules as requirements for membership, but that particular set of strictures may only reflect the pattern of preferences shared by members of that particular (individually self-selected) community independently of those rules.

Certainly, individuals often voluntarily join sects that nominally restrict their behavior in various ways. But a given nominal restriction may only reflect the pre-existing dislikes of the individual adherent. Many people simply do not have libertine preferences, but instead place high value on goals such as security and orderliness; they may also have preferences influenced (in the long run) by their economic circumstances. For example, many of the medieval heresies in Western Europe—that is, black market providers of illegal alternative religious doctrines—rejected material wealth, glorified poverty, and advocated extreme sexual abstinence. However, requiring the (already) poor to reject material wealth is hardly an onerous restriction, and many of the medieval heresies inspired their adherents to reduce the sinful material wealth of the rich by relieving such sinners of their property—another requirement likely to be highly attractive to many poor believers. Similarly, a rule requiring sexual abstinence was potentially a welfare-enhancing *relaxation* of a social norm of producing children, at least from the standpoint of

women, to whom pregnancy and childbirth often had fatal consequences.

Monopoly Religion and Moral Order

A church can be thought of as a firm that markets a pure credence good. Consumers cannot sample the promised life-after-death and use that experience to evaluate the reliability of the church's advertised claims. Consumer acceptance of the church's product depends on their trust in the church's reliability. A firm marketing such an unusual product must carefully protect its brand-name capital by refraining from activities which might be taken as indirect evidence of insincerity or dishonesty. Otherwise, the demand for the services produced might disintegrate altogether.

Therefore, a profit-maximizing firm marketing a pure credence good will avoid activities that might be construed as evidence of opportunistic motives. Elaborate justifications will be developed for payments from consumers to the firm. The very profit-maximizing goal will be disguised for fear of causing consumers to question the validity of the intangible, nonobservable goods available for purchase.

Assume that one religious firm has achieved a pure monopoly, in the sense that competing entrants into the market for religious services are subject to harsh coercive penalties. The monopoly church then behaves like a typical monopolist, setting price by equating marginal cost with marginal revenue in order to extract monopoly rent from consumers.

Historically, churches have tended to extract rent (obtain net revenue) by structuring doctrines in such a way as to require individuals to pay fines "for the good of their souls" to the church. (The church in question typically describes the resulting revenues as destined for various "good works," such as aid to the poor, the glorification of God, or any other goal than the comfort and consumption of church officials.) While the monopoly church might have a number of revenue sources, other things equal, the greater the stringency of moral rules imposed on individual behavior, the greater will be the opportunities for revenue (i.e., fines imposed on sin) and therefore the larger the rent extracted. In a competitive market for religion, sects will face the loss of members to other sects if they implement policies that impose burdens on their present congregation, other things equal. As we noted previously, even some nominally strict requirements may actually impose a zero or negligible burden on particular individuals. However, if one organization holds a legal monopoly

in the market for religion, this process is constrained. Consumers are prevented from shopping for the least-cost set of moral requirements. The welfare of individual consumers may suffer in the short run, but the effective restrictiveness of moral rules will tend to increase. This constraint, in turn, may enhance economic efficiency, if the monopolist has an incentive to produce efficient moral rules.

Further, the very existence of competing religious sects depreciates the credibility of all religious beliefs taken together. If God exists and provides for life after death and other nonobservable attributes, why does He permit the True Church to be challenged by interlopers? The existence of multiple competing sects therefore limits the extent of the market for religion in general. Detailed claims regarding behavioral restrictions and moral rules become more difficult to sell to consumers. Religiosity—the willingness of consumers to actually believe religious claims about the supernatural—represents a problem of the commons from the perspective of religious sects. The larger the number of competing religious sects, the greater the extent of “overgrazing” on the commons becomes. But a monopoly church is the sole owner of the religiosity commons and has an incentive to prevent overgrazing.²

Although religious sects often provide some collective goods to their members, the most basic services they ostensibly provide are purely *private* goods. These services involve the intercession on the believer’s behalf with the divine power. For example, the principal service the Catholic Church purports to provide to its members is salvation, not bingo games or insurance benefits. Salvation is necessarily a purely private good, in which others play no role. (The official doctrine of the Catholic Church has long held that salvation was strictly a function of faith on the part of the individual, and good works only play a role insofar as they represent an expression of that faith.) In pagan religion, the gods were believed to reward their

²Lawrence Iannaccone (1989) offers a quite different analysis of the relationship between moral strictness and the structure of the religious market. He argues that religious sects tend to compete with each other by offering stricter rules and regulations. He postulates that such requirements allow competing sects to more efficiently control free-riding behavior, and therefore increase the attractiveness of membership. However, he models sects solely as providers of various forms of club goods to members, and implicitly assumes that divinity plays no role. But worship of the divine is the main purpose of religious participation. God (or, the gods) have almost always been portrayed by religious sects as omniscient, and extremely powerful when annoyed (and sometimes even omnipotent). In other words, the divine is the ideal monitor against free-riding. Shirking, insincerity, and other forms of malfeasance by members of the religious sect tend to be crimes punishable by divine intervention—either after death or even during life.

faithful servants with good fortune during life. But whether it takes the form of expected reward in an afterlife, or a more tangible reward in the here-and-now, the favor of the deity (or deities) is an ordinary private good (albeit with some extraordinary characteristics).

The case of the provision of moral rules by religion may be an example of the efficient provision of public goods as a by-product of rent seeking by a self-serving interest group. The monopoly church may provide a socially efficient bundle of moral rules as a by-product of its efforts to maximize rent extraction from consumers of religious services.³

Rent Seeking and the Provision of Moral Order

Under some circumstances, a rent-maximizing religious monopoly may tend to benefit from promoting certain kinds of moral rules even though those rules do damage to the overall economy. Such a monopolist may maximize its return in the short run by supplying moral rules that represent public bads in this sense.

For example, by ruling sinful certain forms of ordinary behavior that individuals otherwise find rewarding and that simultaneously serve economic efficiency (e.g., trading in the money economy for profit), a monopoly church might expect to obtain significant rents (in the forms of contributions, penalties, and so on). In such a case, commercial activity would probably decline, and those business activities that continued (although perhaps driven underground) would be a source of fines and other payments to the monopoly church, as transgressors sought forgiveness. Such monopoly rent-maximizing behavior could conceivably result in massive losses to the economy, as efficiency-enhancing forms of individual behavior were systematically reduced. This argument would seem to suggest that monopoly religion might tend to *oversupply* moral rules of an inefficient type.

This short-run model, however, neglects the monopoly religion's long-run incentive to maximize profit. If a monopoly religion benefits directly from economic growth, it will have an incentive to provide moral rules that promote economic efficiency. These direct benefits would take the form of tithes, bequests, and other contributions from believers who view their contributions as normal or even superior goods, which would tend to grow with the wealth of those contributors. Given this assumption, monopoly religious organizations will have a vested interest in economic efficiency, which implies that

³See Cowen, Glazer, and MacMillan (1990) for a recent general discussion of this problem. See Anderson, Rowley, and Tollison (1988) for an historical example.

the provision of economically efficient moral rules will be a wealth-maximizing strategy from the perspective of the rent-extracting monopoly religion.

Compare this vested interest in economically efficient moral strictures with the situation likely to occur under competitive conditions. Individual competing firms will benefit in the short run by providing moral rules their members find most congenial—despite the detrimental long-run effects some such rules could have on economic efficiency. But unlike the monopoly religion, which expects to survive indefinitely and hence has a very long time horizon and an incentive to maximize returns in the long run, competing sects do not expect longevity and thus have short time horizons and an incentive to maximize returns in the short run.

Competing religious sects face a problem of the commons: they will maximize short-run profits by consuming part of the moral capital of society. Any detrimental consequences to the overall economy will tend to be diffuse and borne by all. The short-term advantages of appealing to potential members by offering moral rules that adversely impact the moral constitution of the society will be concentrated and restricted to the opportunistic individual sect in question. In economic terms, moral externalities will fail to become fully internalized. But the monopoly religious firm will be a residual claimant of the gains to society from efficient moral rules and will consequently have the incentive to provide a more nearly optimal mix of moral regulations. The monopoly church will directly benefit from any activity that increases economic growth, because a larger population represents an expansion in its membership, and a richer society will be a source of more lucrative contributions.

Competing religious sects will tend to improve the moral services they provide to consumers by offering moral rules that minimize the moral burden on each adherent. These moral rules will tend to constrain individual behavior less, and justify individuals' pre-existing preference orderings more. Religious sects in a competitive market will evolve toward clubs composed of individuals with similar economic interests and preference orderings. These clubs will provide various forms of insurance to their members, economies of scale in consumption of collective goods, and function as mutual protection societies for people of like preferences. Eventually, some of these clubs will specialize in appealing to individuals with socially inefficient preferences who find themselves rewarded by the peculiar moral strictures offered "for sale" by such organizations. The net effect of this race for the bottom could be that some forms of

economically inefficient, opportunistic individual behavior are encouraged.

A Reinterpretation of the Rise of the West

Our model of religion may have significant implications for economic history. One possible illustration of the economic consequences of monopoly religion involves a reinterpretation of a premier historical conundrum: the rise of Christianity and the fall of Rome.

The triumph of Christianity over paganism in the late Roman Empire is one of the major events in world history and has been extensively studied for centuries. In 313, the Roman Emperor Constantine the Great issued his Edict of Toleration, providing legal protection of Christians from persecution. Over the space of the next 10 years, Constantine gradually made Christianity the official state religion. Finally, in 392, the Emperor Theodosius I made the practice of paganism effectively illegal (Huttmann 1967, p. 216). Christianity became the state monopoly religion of the Roman Empire.

The paganism that preceded Christianity in the Roman Empire was not one religion, but actually dozens—and possibly hundreds. While there was a Roman religion, with its peculiar pantheon of gods and lesser deities within the Empire generally, and even within Rome itself, many other religions were allowed to compete freely for adherents. Aside from occasional legal requirements that nominal respect be paid to the Emperor as a deity, none of the structure of worship of any Roman pagan sect was defended by legislation prior to Constantine the Great's reign (MacMullen 1984, p. 15).

In short, paganism referred to a highly competitive market for religion in the Roman Empire. Entry barriers were essentially nonexistent, and new religions were common. The Roman religion successfully accommodated foreign cults for centuries, by incorporating their deities into the Roman religion as alternative identities for Roman deities.

Pre-Christian paganism in the Roman Empire had the major characteristics the model presented here would have predicted. Most importantly, pagan cults normally had little or nothing to say in the form of ethical restrictions or moral rules. This was particularly true for the Roman religion itself. Religion was presented as a way of interacting with the divine entities who actually ruled the physical world, either to forestall their wrath or to seek their active support for mortal enterprises. As MacMullen (1984, p. 13) explains:

From the divinity all good might be received: foreknowledge, safety in risky doings, good crops, and (attested above all else and espe-

cially signalized in various points of cult) good health. Such were the universal expectations, so far as we can judge, outside of Judeo-Christian circles. The idea that any sort of day-to-day service or perpetual allegiance was owing to divinity had little currency. Neither was there the least uneasiness about the self-interested nature of worship. You only made offerings, or promises of offerings, in order to gain favor from powerful beings.

Pagan Roman religions did not, in general, preach hedonism, or libertinism (although there were some individual religions that did); rather, the answers to moral questions were not assumed to follow from religion. Moral codes were based on tradition, or law, or reason, but were not normally presented as the dictates of the gods. The gods were typically portrayed as quite uninterested in the day-to-day affairs of mere mortals, unless specific mortals had convinced one or more divine entities to pay attention temporarily—usually, it was assumed, because the gods in question thought that they themselves could expect to receive some tangible benefit. This minimal moral constraint is exactly the kind of outcome we would predict to be associated with a highly (if not perfectly) competitive market for religion. Competing sects would be forced to compete by offering potential customers higher quality products, including lenient constraints on personal behavior. Pagan cults attempted to attract new adherents by providing superior entertainment, and by attempting to convince the nonconverted that the cult's deity (or deities) provided better services to followers, and hence were worthy of allegiance (see MacMullen 1980, pp. 18–33).

The primitive Christian church, however, was an exception. Although Christianity in the early centuries stood for a plethora of competing sects and doctrines, the Church that achieved state monopoly status behaved predictably. Moral rules and other regulations on personal conduct were stringent and rigidly enforced. The Christian God was the one and only True God, and the Church condemned all other, competing divinities worshiped by other sects. The Christian God did not offer valuable services to potential adherents, but failure to accept Him was portrayed as leading to horrible torment after death (i.e., hell).

In fact, this intentional exclusivity (or “exclusionism”) was what got the early Christian movement into trouble with Roman authorities. Christians were periodically persecuted, although the persecutions usually involved relatively small numbers of Christians. The only widespread persecutions of Christians occurred during the reign of Diocletian (284–305), and during the first decade of the 4th century preceding the Edict on Toleration. Christians were perse-

cuted because they adamantly refused to acknowledge the divinity of other gods than their own. The Roman authorities had no complaint with any other aspect of Christian religion. Roman toleration of alternative sects did not extend to "atheists," which was the term Roman officials repeatedly used to describe the exclusivity of the Christians (see Chuvin 1990, pp. 17, 21). Furthermore, this attitude toward exclusivity was expressed not merely by anti-atheistic officials, but reflected the view of many Roman citizens. Petitions to the Emperor to initiate persecutions against Christians because of their alleged atheism have survived, and they provide strong evidence that many ordinary people were worried about Christian denial of the validity of other sects and the divinity of other gods (Smith 1976, p. 43).

Roman policy aimed against the exclusivity of Christianity was similar to their policies directed at the adherents of other religions that also claimed monopoly status for the divine, such as the doctrines of Mani and Judaism. Anti-atheist policy, then, was the equivalent of antitrust measures in the market for religion. Sects with a stated ambition to achieve monopoly status, which denied the divinity of the gods worshiped by other sects, were subject to legal sanction.

Naturally, these Roman policies resembled modern antitrust enforcement only imperfectly. There were many differences. Competing sects could not bring lawsuits against the would-be monopolists. There appear to have been no legal penalties for price-fixing, or restraint on trade, in the religious market. (We simply do not know if any significant amount of collusion took place between different sects, although it is well-established that relations between the various religions were ordinarily quite peaceable.) And of course, the Romans did not explicitly speak in terms of "regulating the religious marketplace." Still, the policies had the effect of pro-competition regulation.⁴

The Christian church began to rapidly acquire special governmental privileges with the Edict on Toleration (Milan 313). After Constantine's reign, the establishment of Christianity as a state monopoly religion continued steadily, albeit with occasional brief interrup-

⁴Additionally, some of the instances of anti-Christian persecution appear to have been at least partly motivated by the concern that Christians were potential traitors against the Roman state. There were two reasons for such a concern. The first was that Christians refused to acknowledge the Emperor as a divine entity, to refer to the Emperor as "Lord," or to take oaths on the "Emperor's Genius." The other source of concern about Christian loyalty involved the occasional outbursts of social unrest which were thought to have been led, or at least encouraged, by Christians (Smith [1776] 1976, p. 30).

tions. In 356 worship of images (a practice virtually universal in pagan religions) was technically a capital crime within the Empire, although the law was only sporadically enforced. From about 380 on, organized persecution of non-Christian sects, both by the government and with the governments's tacit or explicit consent, accelerated in frequency and intensity. By 407, "it could be fairly claimed that non-Christians were outlaws at last" (MacMullen 1984, p. 101). The same historian observes:

From Rome in 407 issued a decree to the west, "If any images stand even now in the temples and shrines. . . , they shall be torn from their foundations. . . . The buildings themselves of the temples which are situated in cities or towns shall be vindicated to public use. Altars shall be destroyed in all places." Nothing could be more explicit. It was no longer enough to favor the church, no longer enough to forbid the murkier practices of pagan cults; now anything and everything to do with them must be annihilated [p. 101].

Since the publication of Edward Gibbon's *Decline and Fall of the Roman Empire* (1776–88), some historians have attributed the Roman Empire's demise partially to the rise of Christianity. Gibbon argued that the basically reasonable and down-to-earth Romans were transformed into irrational, other-worldly apocalyptics, and that this change somehow led to the weakening of Rome both militarily and culturally.⁵ Although Gibbon's close friend, Adam Smith, does not discuss Roman paganism in his *Wealth of Nations*, Smith may have had this example of free-market religious competition in mind when he wrote his famous account of the advantages of religious competition in promoting reasonable beliefs and attitudes among the religious.

Gibbon bemoaned the fall of Rome, and claimed that the average person in the Roman Empire at its height was wealthier than his equivalent in late 18th-century England. But on this count, Gibbon was simply wrong. Real per-capita income in 18th-century England was many times higher than it had been in Rome even during the best of times. Moreover, Gibbon appears to have ignored the fact that a sizeable percentage of the Roman population were slaves. He chose to overlook the incessant and destructive civil wars that plagued the Roman Empire throughout its history; Roman government was highly unstable at the imperial level. Finally, Gibbon failed to note the stagnant economy of the Roman Empire.

⁵Gibbon described the early Christian church as dominated by the irrational and those who encouraged social discord, contributing to the breakdown of Roman reason and order. See Gibbon (1974, pp. 346–47) for a summary.

Per-capita income in the Roman Empire appears to have experienced very low growth. According to a recent survey of the problem, the Roman economy was “underdeveloped,” and “the mass of the population lived at or near subsistence level”; the period of the Roman Empire “deserves to be categorized as one of relative stagnation” (Garnsey and Saller 1987, pp. 43, 63).

Compare this record with the history of the Christian West. Although older writers dismissed the Middle Ages as a period of stagnation, modern economic historians generally agree that the period beginning around 1000 until immediately before the Reformation of the mid-16th century—a period in which the market for religious services was dominated by the monopoly Roman Catholic Church—represents a period of sustained economic growth in Western Europe, interrupted by the Black Death of the mid-14th century (Cameron 1989, Jones 1989, and Mokyr 1990).⁶ Thus, the modern period of very high rates of economic growth which began after the Protestant Reformation was a continuation, and acceleration, of a trend that was already 500 years old.

Any measurable influence on the moral constitution of a society exerted by a monopoly religion would presumably require a long period—possibly centuries—to take effect. The Edict of Milan did not lead to any immediate growth spurt in the Roman Empire, and the succeeding five centuries or so appear to have been economically stagnant. We suggest that this delay may have been, in part, a lag in the emergence of a new moral constitution, in addition to having been a period of prolonged civil strife caused by other factors.⁷ Also, the monopoly power of the Roman church was significantly strengthened after about 1050 by the Papal Revolution and related events, and the previous centuries of less-effective monopoly may have muted the effects on the moral constitution (see Berman 1983, Morris 1989).

Our model of religious marketstructure has drastically different implications for economic development than does the well-known alternative argument propounded by Max Weber ([1904–5] 1948)

⁶Cameron (1989, p. 44) summarizes this consensus: “In fact, medieval Europe experienced a flowering of technological creativity and economic dynamism that contrasts strongly with the routine of the ancient Mediterranean world.”

⁷On the other hand, recent economic historians have noted the impressive technological advances that occurred during the “Dark Ages.” These included better and bigger water-wheels, horseshoes, the stirrup, the breast-strap and horse collar, the tandem harnessing of horses, the whippetree, the use of plaster, the use of soap, the use of fossil fuels, the development of chimneys, and others. All of these seemingly mundane but significant innovations originated before 900 (see Mokyr 1990, chap. 3).

and Richard Tawney (1926).⁸ They maintained that the medieval Catholic Church greatly hindered the emergence of capitalism by purportedly teaching that profit-seeking was sinful and by restricting many growth-enhancing activities (e.g., usury); capitalism only emerged after the Protestant Reformation, when a more competitive market for religion allowed for the development of more liberal doctrines regarding business and commerce. This thesis is accepted by a number of recent writers (e.g., Rosenberg and Birdzell 1986).

Our response is that Weber and Tawney were seriously mistaken on several grounds. Two are germane here. First, medieval Europe was not the stagnant swamp they imply, but had a dynamic, growing economy in which capitalism (i.e., monetized exchange) began to develop very early. Second, the Catholic Church did not reject profit-seeking, but only opportunistic behavior (i.e., avarice); ordinary commerce was actively encouraged. Apparent exceptions, like the prohibition against usury, were actually implemented in ways that allowed easy evasion. The monopoly church rather than governments provided many public services vital to the functioning of a market economy, such as legal dispute resolution and charitable institutions.⁹ Given the relevant constraints—both the undeveloped state of the overall economy and the comparatively primitive state of available science and technology—the church behaved in a manner that probably benefited economic growth and development.

Rates of economic growth in the period 1550 to 1850 (and, of course, to the present day) have been greater still than those evidenced during the Middle Ages, and coincide with the aftermath of the Protestant Reformation. Weber and Tawney, and many other historians, have shown a tendency to portray the Reformation as a great expansion in religious freedom and competition between sects. The Catholic Church did face increased competition across Europe, but in many countries the Reformation led to state-sponsored monopoly Protestant churches.

The foregoing account of the history of the later Roman Empire is intended as an interpretive illustration of the economic consequences of the superior provision of moral order by rent-seeking religious organizations protected from competition by governmental monopoly. We have provided only a brief sketch of the incredibly complex events that transpired in the centuries surrounding the rise

⁸Weber was not the first to make such an argument. Adam Smith offered a similar thesis, although it differed in some details (see Anderson 1988). Viner (1978, pp. 159ff.) found several earlier writers who made similar claims, some dating from the late 17th century.

⁹See Gilchrist (1969, pp. 122–40) for a general discussion of these matters.

of Christianity. Certainly, many other factors contributed to the relative stagnation of Rome (with its highly competitive religious marketplace), as compared with the economic dynamism of the Christian West. We merely propose that one important contributing factor was the superior ability of the (coercive) monopoly Christian church to provide an efficient moral constitution for economic growth and development. Unlike pagan religions, once the Christian church achieved a legal monopoly, it took steps to "enforce moral standards and to promote uniformity in theological expression" (Grant 1970, p. xii). Included in the package were those moral standards that enhanced economic growth.

This argument is not intended to imply that the monopoly effect on the moral order is the only, or even the primary, factor in economic development. In fact, it is likely to be more important in relatively less developed settings. As economic development proceeded in the West, institutions designed to minimize social transactions costs have greatly expanded. Formalized, external constraints on individual opportunistic behavior—such as police, law, courts, and capital markets—have become relatively more important, relieving the need for internalized constraints in the form of a moral constitution, whether provided by organized religion or otherwise. In other words, two of the three constraints on individual social behavior discussed by Adam Smith, namely law and human capital incentive effects, have become more important over time, while the third, the internalized monitor, has become less significant.

Conclusion

The consumption choices made by individuals in the market for moral rules collectively generate the moral constitution of society. This moral constitution represents the behavioral infrastructure necessary for an efficient economic order. The sets of moral rules chosen by individuals as guidelines in their personal lives are private goods, but the moral constitution that results from the collective effect of those individual choices is a public good. We have argued that competitive markets for religion may generate negative externalities for society because "profit"-maximizing, competitive religious firms will tend to improve the quality of the moral rules they offer for sale to individual consumers by decreasing the effectiveness of those rules as behavioral constraints. The presence of monopoly in the market for religion may tend to reduce the quality of moral rules from the standpoint of individual believers, but improve the quality of the moral constitution supporting a market society. This result may help

explain why "The European Miracle" (Jones 1981), and other examples of sustained economic growth, have occurred when and where they did throughout history.

References

- Anderson, Gary M. "Mr. Smith and the Preachers: The Economics of Religion in the *Wealth of Nations*." *Journal of Political Economy* 96 (October 1988): pp. 1066–88.
- Anderson, Gary M.; Rowley, Charles K.; and Tollison, Robert D. "Rent Seeking and Human-Exchange." *Journal of Legal Studies* 17 (December 1988): pp. 83–100.
- Becker, Gary S. *The Economic Approach to Human Behavior*. Chicago: University of Chicago Press, 1976.
- Berman, Harold J. *Law and Revolution: The Formation of the Western Legal Tradition*. Cambridge, Mass.: Harvard University Press, 1983.
- Brennan, Geoffrey, and Buchanan, James M. *The Reason of Rules: Constitutional Political Economy*. Cambridge: Cambridge University Press, 1985.
- Cameron, Rondo. *A Concise Economic History of the World: From Paleolithic Times to the Present*. New York: Oxford University Press, 1989.
- Chuvin, Pierre. *A Chronicle of the Last Pagans*. Cambridge, Mass.: Harvard University Press, 1990.
- Cowen, Tyler; Glazer, Amihai; and McMillan, Henry. "Rent Seeking Promotes the Provision of Public Goods." Working Paper. Center for Study of Market Processes, George Mason University, Fairfax, Va., 1990.
- Garnsey, Peter, and Saller, Richard. *The Roman Empire: Economy, Society and Culture*. Berkeley, Calif.: University of California Press, 1987.
- Gauthier, David. *Morals by Agreement*. Oxford: Clarendon Press, 1986.
- Gibbon, Edward. *The Decline and Fall of the Roman Empire: Volume 2*. 1776–88. New York: AMS Press, 1974.
- Gilchrist, John. *The Church and Economic Activity in the Middle Ages*. New York: Macmillan, 1969.
- Grant, Robert M. *Augustus to Constantine: The Rise and Triumph of Christianity in the Roman World*. San Francisco, Calif.: Harper and Row, 1970.
- Huttmann, Maude Aline. *The Establishment of Christianity and the Proscription of Paganism*. New York: AMS Press, 1967.
- Iannaccone, Laurence R. "Why Strict Churches are Strong." Working Paper. Hoover Institution, Stanford, Calif., 1989.
- Jones, Eric L. *The European Miracle*. Cambridge: Cambridge University Press, 1981.
- Jones, Eric L. *Growth Recurring*. New York: Oxford University Press, 1989.
- Lea, Henry C. *Minor Historical Essays*. Philadelphia: University of Pennsylvania Press, 1971.
- Levy, David. "Rational Choice and Morality: Economics and Classical Philosophy." *History of Political Economy* 14 (1982): 1–36.
- MacMullen, Ramsay. *Paganism in the Roman Empire*. New Haven, Conn.: Yale University Press, 1980.
- MacMullen, Ramsay. *Christianizing the Roman Empire, A.D. 100–400*. New Haven, Conn.: Yale University Press, 1984.

- MacMullen, Ramsay. *Corruption and the Decline of Rome*. New Haven, Conn.: Yale University Press, 1988.
- Mokyr, Joel. *The Lever of Riches: Technological Creativity and Economic Progress*. New York: Oxford University Press, 1990.
- Morris, Colin. *The Papal Monarchy: The Western Church from 1050 to 1250*. Oxford: Clarendon Press, 1989.
- North, Douglass C. *Structure and Change in Economic History*. New York: W.W. Norton, 1981.
- Rosenberg, Nathan, and Birdzell, Luther E., Jr. *How the West Grew Rich: The Economic Transformation of the Industrial World*. New York: Basic Books, 1986.
- Smith, Adam. *An Inquiry into the Nature and Causes of the Wealth of Nations*. 1776. Edited by Edwin Cannan. Chicago: University of Chicago Press, 1976.
- Smith, Adam. *The Theory of Moral Sentiments*. 1759. Edited by David D. Raphael and Alec L. Macfie. Indianapolis, Ind.: Liberty Classics, 1982.
- Smith, John Holland. *The Death of Classical Paganism*. New York: Charles Scribner's Sons, 1976.
- Tawney, Richard H. *Religion and the Rise of Capitalism*. New York: Harcourt, Brace and Co., 1926.
- Vanberg, Viktor. *Morality and Economics: De Moribus Est Disputandum*. New Brunswick, N.J.: Transaction Books, 1988.
- Viner, Jacob. *Religious Thought and Economic Society: Four Chapters of an Unfinished Work*. Edited by Jacques Melitz and Donald Winch. Durham, N.C.: Duke University Press, 1978.
- Weber, Max. *The Protestant Ethic and the Spirit of Capitalism*. 1904-5. New York: Free Press, 1948.

IN DEFENSE OF THE CREDIT BUREAU

Daniel B. Klein and Jason Richner

Introduction

In the past year Congress considered more than 30 resolutions concerning credit bureaus. The bills were designed to expand control of credit bureaus. The House Subcommittee on Consumer Affairs and Coinage heard hundreds of stories from consumers claiming to have lost their jobs or been denied credit because of credit bureau mistakes.

The big credit bureaus—TRW Inc., Equifax Inc., and Trans Union Credit Information—realized that legislation was unavoidable and endorsed a bill by Rep. Esteban Torres (D-Calif). The Torres bill (“The Consumer Credit Reporting Reform Act”) was the most likely to pass. The bill contained a “preemption provision” that would have given the federal law preemption over state regulation of the credit reporting industry. On September 24, 1992, after several hours of debate, the U.S. House of Representatives rebuffed an attempt by the bill’s sponsors to remove the preemption provision (the provision that ultimately led to an endorsement by the credit reporting industry) by a 207 to 203 vote. Rather than go forward with the bill that retained the preemption clause, the sponsors pulled the bill, with the 49 measures it contained. Although the industry may have gained some respite, the 50 individual states may now act more aggressively to pass their own legislation, and Congress is renewing its efforts at present.

The result for consumers, supposedly represented by Consumer’s Union and Ralph Nader’s U.S. Public Interest Research Group (PIRG), may be more restricted credit, less attractive terms for credit (such as higher interest rates), and higher prices in general. As a

Cato Journal, Vol. 12, No. 2 (1992). Copyright © Cato Institute. All rights reserved.

Daniel B. Klein is Assistant Professor of Economics at the University of California, Irvine, and Jason Richner is a free-lance writer.

consequence of the crusade against credit bureaus, many consumers will be unable to get credit.

When Torres introduced the bill, he said the current law allows the private lives of consumers to be "an open book." "Sensitive personal and financial data is bought and sold without regard to privacy; consumer reports are riddled with errors, resulting in denial of credit, insurance and employment" (Herubin 1991, p. C1). To get an idea about what other lawmakers have been thinking, we need only read the comments of Rep. Bernie Sanders (I-Vt.) During the markup of the Torres bill, Sanders said that the subcommittee should not be debating amendments to the bill, but should be talking about whether credit bureaus should exist at all. (*Communicator*, April 1992, p. 6). Congresswoman Maxine Waters (D-Calif.) added, "It was outrageous for members to feel compelled to protect the credit reporting industry" (p. 6). The assault on the credit reporting industry may be just beginning.

In response to criticism from consumer groups, TRW Inc. decided in 1992 that it would provide consumers with a free copy of their credit report. The announcement came soon after TRW admitted to having made a mistake with records from Norwich, Vt. All 1,400 taxpayers were branded as delinquents when a part-time clerk went to Norwich to collect names of delinquent taxpayers. She mistakenly wrote down the names of everyone who paid their taxes. TRW's new policy put considerable pressure on its two rivals, Equifax and Trans Union, to follow suit. Equifax responded by saying that it expects to cut the price to consumers of obtaining a copy of their credit report. Trans Union was not as quick to give in, saying that providing free reports is an "extremely expensive solution." "We will continue to implement real solutions" said Trans Union, "rather than 'public relations fixes'" (Miller 1991b, p. B5). Since further legislation might hamstring the industry, credit reporting agencies seek to appease so-called consumer groups.

Critics claim that the industry has undergone so many changes in the last 20 years that existing legislation is obsolete. The industry seems resigned to the view that "it must support new legislation to win back the public's favor" (Miller 1991a, p. B1). The purpose of this paper is to show the social benefits of credit reporting and the foolishness of the campaign for tighter regulation. The practices and procedures of the credit reporting industry are examined and we respond to the concerns of Consumer's Union, PIRG, and others who favor aggressive legislation.

The Social Function of Credit Bureaus

Credit enables millions of consumers to buy millions of dollars worth of goods every day. Yet the value of credit extends beyond

the fact that using credit is safer and more convenient than carrying large sums of cash. Credit gives the consumer the ability to arrange a consumption plan better suited to his individual earning pattern. He can finance real-estate and consumer durables to smooth out his consumption over his life. Although many consumers do themselves a disservice by over-using credit, the fact that millions of consumers use credit everyday is strong evidence that it is a useful tool.

It is crucial to realize that consumers can get credit precisely because credit granters can identify which consumers are likely to pay their bills. Without credit bureaus, businesses would have a tough time accumulating the payment histories of individual consumers and would not give credit, except in special circumstances. Punishing a consumer who defaults would be an expensive and time consuming process. Historically, when credit reporting is absent, so is consumer credit (Cole 1988, pp. 184–185). When credit reporting is in place, consumers have an extra incentive to pay their bills. They are eager to keep their credit report clean, for otherwise they may lose the benefits of credit. Besides creditors, apartment managers and prospective employers sometimes consult credit records. By enhancing accountability, credit bureaus help turn consumers into responsible individuals.

The ability of creditors to receive information about individual consumers has far-reaching benefits. If credit granters did not have the knowledge gathered by credit reporting agencies, then creditors would be forced to make up for the losses caused by defaulters. Prices would have to increase, harming all consumers. Faithful bill-paying consumers would not be rewarded for their good behavior.

According to the President of Associated Credit Bureaus, Walter Kurth (1991, p. 2), “The mission of the consumer reporting industry is to serve as an objective third-party provider of information to the companies and consumers involved in credit transactions. Our members are libraries that make it possible for credit granters to provide consumers with the credit opportunities they seek.” Additionally Kurth (1991, p. 2) points out that “150 million credit active Americans benefit from the opportunities made possible by credit extension.” Almost one billion credit cards are in use in the United States and millions of Americans finance homes by a mortgage. Credit is made available to anyone in the United States who can show that he or she has the ability to pay loans as agreed. “American citizens’s access to housing, cars, televisions, clothing, household appliances, travel, education, and medicine is understandably the envy of the world” (Kurth 1991, p. 3).

The Credit Bureau as a Mechanism of Social Control

Adam Smith said the following in 1763 in a lecture entitled, "The Influence of Commerce on Manners,"

Of all the nations in Europe, the Dutch, the *most commercial*, are the most faithful of their word. The English are more so than the Scotch, but much more inferior to the Dutch. . . This is not at all to be imputed to the national character, as some pretend; there is no natural reason why an Englishman or a Scotchman should not be as punctual in performing agreements as a Dutchman. It is far more reducible to self-interest, . . . (which) is as deeply implanted in an Englishman as in a Dutchman. A dealer is afraid of losing his character, and is scrupulous in observing every engagement. When a person makes perhaps twenty contracts a day, he cannot gain so much by endeavoring to impose on his neighbors, as the very appearance of a cheat would make him lose. When people seldom deal with one another, we find that they are somewhat disposed to cheat, because they can gain more by a smart trick than they lose by the injury which it does their character [Smith (1763) 1964, pp. 253–54; italics added].

Credit bureaus help to turn consumers into Dutchmen. The credit bureau provides the links between the relationships the consumer has with various creditors. This linkage gives the consumer a reputation, or in Smith's words, a "character." The possibility of black marks looms large if the consumer attempts any "smart tricks." One of the first threats that a creditor makes to a delinquent is to report the matter to a bureau (Cole 1988, p. 246).

In a small community members know each other and exert social control through informal means such as gossip. Reputations are made, stored, and transmitted by gossip. The term gossip is used by anthropologists and sociologists to encompass the idea of community information transmission achieved through group meetings, correspondence, local newspapers, leaflets, word of mouth, and so on (Merry 1984, pp. 275–79). In contrast, people in a larger society are mainly strangers and cannot rely on informal reputational means. Man stands, as Adam Smith ([1776] 1937, p. 14) put it, "at all times in need of the cooperation and assistance of great multitudes, while his whole life is scarce sufficient to gain the friendship of a few persons." Credit reporting agencies are a formalized system of gossip that help overcome the gaps in knowledge about particular individual consumers. Without the credit reporting system that makes possible consumer reputations, the consumer could simply default at one business and then apply for credit at another. Foreseeing the likeli-

hood of default, businesses would not give credit. When businesses share their information through a credit bureau, promise-keeping is encouraged through reputational enforcement by credit granters.

Researchers have shown that cooperation often prevails in situations without a policeman wielding a night-stick over deviant behavior. There are two main reasons for this: First, because norms are internalized; people develop the habit and taste for cooperating (Elsater 1989). Second, even though real community interaction is haphazard, information concerning performance travels through the community. People gossip.

American society, however, is filled with countless needs and wants that are difficult to handle and control. Gossip alone cannot handle the requirements for maintaining social order. One could simply leave town to make a new name and reputation for oneself. So, how could proper conduct be promulgated in the Hayekian vision of an extended, anonymous, and flourishing society (Hayek 1973, 1976, 1979)? Institutionalized gossip (such as credit bureaus) can help create the rewards in an extended society the way gossip does in the small community.

By drawing a parallel between credit reporting and gossip we run the risk of playing into the hands of the critics of credit reporting. Although gossip does serve useful functions, as described by Merry, it is often malicious and intrusive. That is precisely how critics portray credit reporting. For example, the Consumers Union study is entitled "What Are They Saying about Me?" (29 April 1991). We must emphasize *in this respect credit reporting is remarkably unlike gossip*. Gossip may be unsubstantiated and anecdotal, it may be exaggerated and peculiar, it may be malicious and manipulative. In contrast, a credit report is the result of professional, standardized, and impersonal procedures.

Contrary to the impression given by the so-called consumer groups, there is a great deal of simple information that credit bureaus do not know about individual consumers. They do not know

- where one went to school and what degree one has,
- an individual's driving record,
- whether one has pets,
- who one's relatives are, except sometimes the spouse,
- one's religious, political, or cultural views.

Although the credit bureau knows that a consumer has a mortgage and a credit card, and may know his place of employment, its knowledge includes a very small part of the basic facts about a consumer's existence, facts that a casual acquaintance might know.

A notorious feature of all social control mechanisms is that individual accountability must be weighed against privacy. A social control mechanism attempts to punish wrong-doers; first, it must *determine* who did wrong and what wrong was done. Questions must be asked, and information must be passed. *All mechanisms of social control collide with privacy.*

Again, gossip is a useful tool for spreading knowledge about the character of individual persons. For instance, a baby-sitter who abused children would not be able to maintain a favorable image for very long as news about his practices spread through town. To guarantee by law that this baby-sitter have an absolute right to privacy (in effect, banning conversation about him) would be both an injury to the public welfare and an egregious infringement on the liberty of *others*. The welfare of those being abused and the liberties of those engaging in communication must also be recognized.

But this is not to throw out the idea of privacy, it is to admit that trade-offs must be studied and judged with care. Viewed against the performance of other mechanisms of social control, one finds that the credit bureau is remarkably respectful of privacy. Compared to the sensational tactics of the press, the entrapment and wiretapping practiced by police, the taintedness of gossip, and, as illustrated by the Clarence Thomas hearings, the disclosure of public testimony, credit reporting must be deemed a precise and unintrusive means of social control. It deals with standardized, automated information, and utilizes only the most relevant information and transmits it to only the most relevant parties. Credit bureaus are discreet and impersonal when gathering and storing information precisely because it is in their best interest to act this way. They serve an important social purpose and clearly are less intrusive than other social control mechanisms that we accept without a moment's thought.

What Do Credit Bureaus Do and How Do They Work?

Most of the 1,000 automated credit bureaus operate on one of the three on-line data processing systems: Trans Union, TRW, or Equifax. But nearly 500 of these automated credit reporting agencies are owned either by private independent parties or a non-profit merchant association, purchasing information from one of the three giants (Kurth 1991, p. 1).

A credit bureau provides a report of a consumer's credit history. The reports contain a list of open credit lines (such as a Mastercard account), of outstanding credit balances (such as how much is owed

IN DEFENSE OF THE CREDIT BUREAU

on a car), and a history of payments. Credit reports detail delinquent payments and note any liens against a consumer. The reports contain the individual's address and Social Security number. This information is gathered from creditors, who supply their records either directly to the three giants or indirectly to local bureaus. Trans Union, TRW, and Equifax enhance their database by obtaining information from county, state, and federal repositories relating to bankruptcies, court judgments, and liens.

The Fair Credit Reporting Act (1971, p. 2), which regulates the activities of all consumer reporting agencies, defines a credit reporting agency as

any person which, for monetary fees, dues, or on a cooperative non-profit basis, regularly engages in whole or in part in the practice of assembling or evaluating consumer credit information or other information on consumers for the purpose of furnishing consumer reports to third parties.

The Fair Credit Reporting Act requires each agency to keep accurate records and investigate consumer complaints about information on reports. If a contested item is confirmed by the credit granter, the consumer is entitled to file a brief statement setting forth his version of the story. When a consumer is denied credit based on information in an agency's report, the consumer is entitled to receive a free copy of the report.

It is crucial to realize that answering consumer complaints is an inherent part of the credit reporting industry. There are standard practices the industry follows based on both legal necessity and sound management. It is in the best interest of the industry to maintain accurate records and be responsive to consumer complaints. Input from consumers is precisely the way that credit bureaus check the reliability of their procedures and improve the quality of their product.

The third parties, who receive reports, are usually credit granters (banks and retail businesses), but employers, landlords, and insurance companies also have an interest in the reports. Businesses which subscribe to a credit bureau pay a monthly fee to the credit bureaus and they pay a unit price per report. "Each bureau establishes its own price" (Cole 1988, p. 196). Credit granters that choose not to subscribe have to pay a great deal more for reports and go through more trouble to get them.

Most credit granters give the information concerning their consumers to at least one credit reporting agency. "The contract credit granters sign to become bureau subscribers specifies that they provide their ledger experience promptly and accurately" (Cole 1988, p. 190).

There are two types of firms that contribute computerized data from their accounts receivable ledgers: (1) businesses that subscribe to a bureau in the same geographic area and use the bureau to obtain consumer credit reports, and (2) national firms, such as Sears, Texaco, and VISA, which may buy reports directly from one or more of the giants. Most local and national credit granters give their data to the three giants. These three agencies contain roughly the same information.

TRW publicizes the fact that it maintains credit information on more than 133 million consumers, and that more than 8,000 businesses regularly contribute their accounts receivable data to TRW. Trans Union says it maintains 170 million credit files, covering about 80 percent of all U.S. households (Cole 1988, p. 202).

The consumer has several reasons other than the preservation of his credit record for responsibly paying his department store credit card bill: continuing relations with the business, the fear of legal recourse by the business or a designated collection agency, or just a sense of honesty. The defaulting consumer could, however, just begin using cash or a bank card at the business. And as for legal action, collection agencies find that beyond nasty letters, it is not practical to go to court for receivables below a certain amount. Thus, for small amounts, a consumer pays his bill to maintain a clean conscience and a clean credit report.

Who Is at Fault for Errors in Credit Reports?

The frequency of errors committed by credit bureaus is the main focus of the critic's attacks. The credit bureaus are held solely responsible for errors, but it is imperative to realize that many forces are involved in the credit reporting process.

One cause of errors that is often overlooked is the United States Postal Service (USPS). The failure of the USPS to forward mail promptly and properly often leads to problems. When a bill is not forwarded the inevitable result is that the consumer will be reported delinquent. The credit bureau is not involved in any way; it only records the information as it is received.

Credit granters are a bigger part of the credit reporting process. Credit granters turn their information over to the credit bureaus. Any errors in the information sent unavoidably becomes part of the credit bureau. The source of credit transactions begins with the consumer and the credit granter. The credit bureau merely records the information as it is received.

Consumers also play a vital role in maintaining accurate reports, and consumers can make mistakes as well. Bills can be misplaced, personal checks may not be recorded and memory fails, out-going mail can fail to find its way to the mailbox. If a consumer fails to pay the amount owed (purposely or otherwise), a black mark will eventually end up on his report.

Errors can also arise from public records consulted in compiling credit records. Any misinformation in these records will turn up on credit reports. Because credit bureaus are the most visible and centralized component of the credit reporting process, they bear the brunt of complaints. Rightous consumer advocates point the accusing finger at the mysterious credit bureau because that pantomime creates the greatest sensation. They rarely point the finger at the thousands of other institutions involved in the process. Although credit bureaus do occasionally err, for the most part they merely record the information as it is received.

The Error Correction Procedure

A major criticism of the credit bureaus has been that they are slow or even refuse to correct mistakes. Consumers and critics alike must realize that a credit bureau cannot change a consumer's record on the basis of a complaint made over the phone. If the process were this simple, every defaulter would crowd the phone lines with stories of malpractice. When a consumer makes a complaint that a credit report contains false information the credit bureau seeks to verify immediately their information with the credit granter. According to Kurth, reverification usually takes place within two weeks. Additionally, although the law does not require it, it is industry practice to delete derogatory information in 30 days or to explain to the consumer in the next two weeks why it has not been deleted. Credit bureaus have procedures to respond to inquires, and laws are not necessary to mandate responsive behavior. The chain of communication involved—from the consumer to the credit bureau to the credit granter and back—makes any blanket requirement of less than four weeks unfeasible. Mail delivery on three pieces of mail may take two weeks alone. Finally, when a credit granter does verify a delinquency, the consumer can make a statement saying why he thinks the credit granter is at fault, and the statement is included in the report.

Credit granters often give the customer some leeway when making payments. If a customer has shown a good payment record the credit granter often extends the amount of time allowed for the consumer

to pay. In other words a consumer may be late, but not reported as delinquent. However, this lack of uniformity also leads to errors. Kurth emphasizes that the industry needs to work together on the computer tape format and the reporting delinquency period. Correcting errors is a major concern of, and in the interest of, all parties involved. Of the many players that take part in this process, the credit granter and the consumer must share some of the responsibility for accuracy.

The Consumers Union Study: "What Are They Saying About Me?"

According to a 1979 study by the Credit Research Center at Purdue University, most consumers have at best a dim awareness of what credit bureaus do (Dunkelberg 1979, p. 6). This point should be kept in mind when thinking about the informal surveys conducted by Consumers Union and PIRG.

In 1991, Consumers Union (CU) conducted a survey of 161 credit reports for the three credit-reporting giants. CU conducted its survey among *its own employees and acquaintances*. Rather than choose a random sample, CU chose 57 consumers who, were "either employed by Consumers Union (CU) or were acquaintances of CU employees" (p. i). The respondents were asked to describe "the accuracy, completeness, readability, and confidentiality" (p. i) of the reports. We hazard to suggest that the 57 employees and their friends were fully aware that CU was hoping to find errors in the data. Furthermore CU says, "CU worked individually with participants who had questions in interpreting their reports in filling out the questionnaire" (p. 1).

CU determined that "the credit reports reviewed were riddled with errors" (p. i).¹ In the "Executive Summary" of the report, CU (p. i) lists the following as the three most central findings on inaccuracy:

1. 48 percent of all the 161 reports reviewed contained inaccurate information. Many reports contained more than one inaccuracy.

¹It is interesting to compare credit bureaus to other major social institutions in the United States in terms of errors and accuracy to put things in perspective. According to Associated Press, "The food-stamp program cost an extra \$1 billion in 1990 because of mistakes, fraud and checks lost in the mail, Agricultural Department documents reveal." Additionally, "Just \$84 million of that was recovered, say officials at the USDA's Food and Nutrition Service" (*Orange County Register*, 15 December 1991, p. A40).

2. 19 percent of all the reports reviewed contained a 'major' inaccuracy—i.e., one that could adversely affect a consumer's eligibility for credit.
3. One CU employee was actually denied credit during the course of the study because of inaccurate adverse information in her credit report.

These claims deserve examination. The 48 percent of 161 reports that contain an "inaccuracy" are claims made by the participants. These claims may be true or false. For the second finding, a "major" inaccuracy is defined as one that could adversely affect a consumer's eligibility for credit. We are not told what exactly CU means by "could adversely affect credit" eligibility. What is the exact nature of the "inaccurate information"? Is it wrong addresses, wrong Social Security numbers, or misspellings? Is it false reporting of a delinquent payment or mere incompleteness? We are never told. CU gives the following example, "If a credit report showed a 60 day late payment, and the participant identified that as inaccurate because he or she made the payment on time, CU included the 60 day delinquency as a 'major inaccuracy' for purposes of this report" (p. 4). Do the respondents have a clear recollection of when they posted their payments, or do creditors, who employ a staff to record payments received, have the reliable information? CU does not follow up on and substantiate the claims of the respondents.

For the third finding, even if taken at face value, it is important to realize that credit granters deny credit for a wide variety of reasons. Furthermore, the alleged error may not be the fault of the credit bureau. CU does not indicate the source of the error. As we have discussed, the credit reporting agency is not the only source of errors.

Another example of a CU staffer's misfortune occurred when a "department store failed to record her new address" (p. i). CU says the bill was sent to her old residence and the post office never forwarded it. The consumer is responsible for notifying all necessary parties about the change of address. How can we be certain that she has? This is simply a mail forwarding problem. The fault lies with either the postal service, the department store, or the consumer, not the credit bureau.

Another case of misfortune is provided by CU in a footnote. One of the CU staffers "had closed several credit card accounts last year by cutting up the credit cards and returning them to the issuers" (p. 3). She was concerned because TRW reports do not clearly indicate when an account is open or closed. But does cutting up your credit cards constitute closing your account? The credit card industry is

plagued by scams, and they must follow careful procedures concerning the opening and closing of accounts.

CU gives another example where an attorney “was delinquent in paying a \$19 balance on a department store credit card that she cut up and returned to the department store 5 years earlier” (p. 5). Did her account have a zero balance when she returned her card? CU never investigated the matter. Also, even if the department store is at fault, there is no cause for blaming the credit bureau.

CU goes on to divide the major inaccuracies into two categories: crossover of consumer files—due to identical names, records are sometimes entered to the wrong file—and “information indicating that the participant was late in making payments when, in fact, the participant made the payments on time or was not obligated to make any payments” (p. 5). Yet, not a single major inaccuracy described is clearly the fault of a credit reporting agency.

Another example involved a staffer who was mistakenly billed for medical services she never received. “Although she cleared up the problem with her doctor—who acknowledged the mistake—her doctor never cleared up the problem with the credit bureau” (p. 5). If a doctor never contacts a credit reporting agency with one of *his* mistakes, one can hardly blame the credit bureau. In its discussion, CU fails to give a single clear case of credit bureau malpractice.

Participants were also asked to judge “whether their reports listed all their current credit accounts” (p. 6). The participants reported that 49 percent of their reports were not complete. Sometimes mortgage loans were absent from reports. Again the question of who is at fault should be posed. Credit granters are responsible for reporting new accounts to credit reporting agencies. Banks are responsible for reporting mortgage loans. Furthermore, credit bureaus do not claim to have a complete record of the individual’s credit activity.

At one point CU says, “many participants also indicated that their reports omitted past credit accounts that had been fully repaid—information that could enhance a consumer’s eligibility for future credit” (p. 6). Credit reporting agencies do not list every account that has ever been held by a consumer. After a certain period of time, information is often obsolete or irrelevant to the credit granter. By updating reports, credit reporting agencies are not bound to punish repeatedly a consumer for an old delinquency.

Finally, CU examined whether participants found their credit reports readable. A multiple choice format was used with the choices representing various degrees of readability. CU offered choices ranging from “Easy to read and understand” to “Extreme torture.” From the results CU concludes “that many participants found the credit

reports difficult to read and understand" (p. 7). Yet, CU's own graph shows that the "Moderately easy to read and understand" category tallied the most responses. But more significantly, credit reports are designed to transmit information concisely to professional parties with a legitimate interest. It is a specialized trade that necessarily employs a specialized format.

In sum, the Consumers Union study is replete with flaws (both semantic and statistical), not the least of which is CU's decision to involve only their employees and friends and to assist them in filling out the survey. The primary goal of the Consumers Union seems to be to inflame opinion against the credit reporting industry, which enables millions of credit active consumers to make purchases they need and enjoy.

PIRG's Study: "How the Credit Bureau Ruined My Life"

Ralph Nader's PIRG also indicts the credit bureau in a study documenting consumer complaints. The study, released on June 12, 1990, examines individual complaints against credit bureaus.

The initial claim PIRG makes is that "credit reporting agencies fail to correct mistakes and fail to properly guard personal credit information" (Mierzewski 1990, p. 1). This study describes "five case studies examining the worst problems found" (p. 1). PIRG claims that approximately 43 percent of credit reports contain serious errors. PIRG wishes to have the Fair Credit Reporting Act, which it calls a "dinosaur" (p. 1), revised because it fails to protect consumers from a variety of significant problems. The number of errors and the difficulty consumers have correcting them seem to be the major thrust of PIRG's document. PIRG relies on quotes from consumers to support its cause for regulation, such as, "To state that the credit bureau ruined my life would not be an understatement" (p. 2), the malapropism PIRG borrowed from when titling its report.

PIRG lists individual complaints representing the most severe horror stories they could find. One consumer says, "*no one I have contacted has any idea how to cure this problem*" (p. 5). PIRG does not inform us whom the person contacted. Another complaint reads, "Early part of last year I was denied because of wrongful report. . . . I was furious as this was to be corrected the first time. . . . *There are exactly seven accounts that do not belong on here and I want them removed for good*" (p. 5). This is another case of the consumer's word against the agency's. Another consumer claims that a credit reporting agency "refused to correct its errors" (p. 5). All of these

claims have one thing in common: they are unsubstantiated and the reader is asked to take them at face value.

"Associated Credit Bureaus (ACB), the trade association for the three repositories, argues that in 1988 'only' about 9 million consumers requested a look at their files" (p. 9). Of these 9 million, 3 million inquiries resulted in an amendment to the report. Three million of 9 million is one-third. On this basis, PIRG suggests that of the 400 million credit reports kept annually, one-third, or 133 million are faulty. The original sample of 9 million consumers, however, hardly constitutes a random sample. The 9 million consumers who paid money to look at their reports probably had reason to do so. Although many consumers undoubtedly bought their reports to satisfy their curiosity, the majority of these 9 million most likely believed that their reports contained controversial information. Thus, to extrapolate this non-random finding over the entire population of 400 million reports is irresponsible.

PIRG recommends that consumers should be able to obtain a free copy of their report every year—a recommendation that TRW has already given in to. Credit reporting agencies charge very little for credit reports as it is; and by law they provide a free report upon request to those denied credit because of adverse information on credit reports.

Another section of PIRG's report concerns "Incorrect Information Not Corrected By All Repositories" (p. 4). One consumer exclaims, "I am appalled to know that there are three universal credit reporting agencies who exchange negative information, but do not share beneficial resolutions" (p. 4). Yet, current anti-trust law forbids credit reporting agencies from exchanging corrected data. PIRG admits that "this is an issue" (p. 4). Furthermore, PIRG concludes "that consumers like Jane Doe are the ones who suffer from bad credit reporting" (p. 4). PIRG fails to recognize that errors, after all, also hurt credit bureaus and the credit granters which rely on credit reports.

The follow up to PIRG's original study is entitled "Don't Call; Don't Write; We Don't Care" (Mierzwinski 1991). PIRG is nothing if not bold. Similar to the language in the first study, PIRG says, "the consumers in this study are trapped in the nightmare on credit street. As this report shows, for many of them the nightmare never ends" (Mierzwinski 1991, p. 1)."

The study profiles 155 "credit report complaints" (p. 2). In numbering its major points, PIRG left out the integer 5—the points are numbered 1 through 4, 6 through 9. Another error on the same page is significant. A bar graph shows the number of "Times Consumer

Contacted Credit Bureau By Phone Or Mail Or Visit.” Sixty-three percent of consumers are shown to have contacted a credit bureau 5 or more times, while 27 percent of consumers contacted a credit bureau less than 5 times. No one at PIRG noticed that 63 percent and 27 percent add up to only 90 percent. The second figure should have been labeled 37 percent, and from the calibration of the axis one can see that they meant 37 percent. These errors found in the 1991 PIRG report show that we all make mistakes. In operations like the credit department of a store, the U.S. Postal Service, or a credit bureau, hundreds of people are employed full-time to record and transmit trillions of bits of information. PIRG has demonstrated that it is unreasonable to demand infallibility.

The study claims, “63 percent of consumers had contacted the credit bureau 5 times or more before contacting the government” (p. 2). PIRG claims that “this is a conservative number, because U.S. PIRG rounded all contacts greater than 5 and converted ‘numerous’ or ‘many’ to 5” (p. 11). By automatically converting “numerous” or “many” to 5, PIRG is hardly being a generous opponent.

Finally, PIRG asks (pp. 3–4), “Why should a consumer who has never done business with a firm be saddled with the hassle, and it is a hassle, of removing someone else’s negative information from his or her credit report?” This question has a simple answer. It is vital for a consumer to contact the firm because that is the only way errors can be corrected. A credit bureau can not correct errors it does not know about. Usually, errors exist because the credit bureau has been supplied misinformation. There is no way for a credit bureau to recognize a discrepancy without the help of the consumer.

The central issue of the debate is the accuracy of the credit reports. Both the Consumers Union and Nader’s PIRG arrive at their estimates by rigging the procedure. The two so-called consumer groups use outlandish and sensationalized language to push their crusade.

The Arthur Anderson Study

Juxtaposed to the histrionics is the study by Arthur Anderson & Co., an internationally recognized accounting and consulting firm commissioned by Associated Credit Bureaus in June 1991 to discover how frequently credit errors cause negative credit decisions (Arthur Anderson 1992). Anderson & Co. studied a random sample of 15,703 consumers that had been denied credit based on the credit granter’s scoring system that used credit report information. Anderson & Co. determined that from this sample, 1,223 consumers (or 7.7 percent) requested their credit report from the issuing credit

bureau. Of this group, 304 consumers disputed the information on the report. Each dispute was reinvestigated and then all 304 resulting reports (some revised and corrected, others simply confirmed) were returned to the credit granters that originally denied credit to the consumer. The credit granter was then asked to reevaluate the credit application. As of the date of the study, 267 of the 304 cases had been reevaluated. In only 36 of these cases was the original credit decision reversed to a positive decision.

The results are telling, but a definitive and exact measure of the frequency of errors cannot be inferred. Consider two cases, the first is the worst case scenario, the second a reasonable guess scenario.

Case 1: Assume that consumers with errors on their report are *no more* likely, after being denied credit, *to request their credit report or to contest information* than consumers without errors. In this case, the 267 reevaluated reports would constitute a random sample of those denied credit, and we would simply divide the 36 decision reversals by 267 to arrive at the estimate that 13 percent of credit rejections result from errors on credit reports. Even though this figure is based on wildly implausible assumptions slanted against the credit bureau, this figure makes the estimates of error frequency of Consumers Union and U.S. PIRG look excessive.

Case 2: A responsible estimate from the Anderson information would of course take into account the fact that someone with an error on his report is more likely to request the report and more likely to contest the information. Suppose that a person with an error were three times as likely as a person without an error to request a report, and four times more likely to contest information. Then the percent of credit rejections due to errors on reports would fall to about one percent. The assumptions here seem like reasonable guesses.

In brief, the study conducted by Arthur Anderson & Co. gives lie to the dramatic claims about error frequency promoted by the so-called consumer groups.

The Issue of Marketing Lists

The issue involving mailing lists is one that feeds on the paranoia of consumers. "What Are They Saying About Me?" is a typical example of tapping the fears of consumers rather than advancing understanding. Merchandisers buy mailing lists from the three largest credit reporting agencies. They buy these lists hoping to find consumers that might have an interest in what they are trying to sell.

It is important to understand how TRW sells mailing lists to marketers. A marketer requests a list of names of individuals that exhibit

characteristics specified by the marketer. For example, the characteristics may be estimated income, zip code of residence, or positive payment history. The marketer sees *only the final list of names and addresses that is compiled, not the information in the files of those individuals*. TRW does not wholesale its consumer information—to do so would be to share freely its stock in trade (Bibas 1992).

Some people undoubtedly regard mail-order catalogs as a nuisance, but if more people did the practice would probably stop. Yet, it is not only businesses that benefit from these transactions. When consumers shop at home by mail-order they save time. Furthermore, they stay off roads and thereby diminish congestion, consumption of gasoline, and auto emissions. The individual consumer is served because catalogs are sent that cater to his or her taste. Without the use of marketing lists, businesses would be reduced to sending out catalogs more randomly or not at all. Consumers Union acknowledges in *Consumer Reports* (October 1991, p. 644) that “readers said they enjoy receiving catalogs from different companies.” Also, “when catalogers narrow the market, they reduce mailing costs, raise the percentage of sales, and make a higher profit. That knowledge has spawned some 10,000 specialty catalogs, selling everything from apples or automobiles to waders or wine. There are catalogs for tall people, short people, left-handed people, people with physical limitations, and people who wear uniforms” (p. 643).

Again, credit reporting agencies are not the only player involved in marketing lists. It is far more reasonable to think that unhappiness with junk mail points to a failure of the postal service. The United States Postal Service (USPS) should perhaps be blamed for not having developed a way for the individual to forestall unsolicited lesser-class commercial mail. If the USPS were more responsive to recipient demands, it could arrange for pre-approval to receive such mail. Under this arrangement individuals would have to think seriously about the information/opportunities versus nuisance of direct mail, and then decide. The vast majority would probably choose to accept such mail. After having been able to make a choice, a consumer would feel less indignant about finding his mailbox stuffed with catalogs.

Another route for individual approval is better awareness that one's name will be used in mailing lists. The suggestion of critics (Consumers Union, PIRG, and others) that marketing firms receive individual approval for every name is obviously too costly. This mandate would, in effect, eliminate direct mail altogether.

Now it should be asked which principles will prevail in the matter of mailing lists. Are individuals to own information about them-

selves? How could this principle find boundaries that avoid absurdities? How would the media report the news? When a business writes down what a person buys to keep track of inventory, is this a violation of privacy? A mailing list is just ink on paper, both of which are the property of the marketing firm.

Conclusion

Credit bureaus are a service to all consumers who use credit responsibly. The agencies serve as a hub of information, efficiently transmitting information to its members. At the same time the credit bureau creates incentives for consumers to alter their behavior. In other words, credit bureaus reward moral actions. For maintaining accountability and honesty in a vast society like the United States, the fundamental principle behind the credit bureau can be of great importance. Essentially, credit reporting agencies play the role that can no longer be played by community reputation. And as a mechanism of social control, credit bureaus are remarkably discreet. There are many institutions involved in the credit reporting process, all of which can be at fault for errors that exist in credit reports. Credit bureaus try to resolve discrepancies, but it is necessary that we recognize credit granters and consumers as important factors in this process.

The Consumers Union and PIRG's reports are rife with irresponsible methods and blatant sensationalism. Consumers Union based all of its findings on the reports of 57 of their own employees and friends. Even with this obviously biased sample, CU fails to point out one error that is clearly the fault of a credit reporting agency. CU does nothing to prove that the mistakes made by credit reporting agencies are above a reasonable level. PIRG's reasoning and statistics also are dubious and sensationalized. In addition to making numerous mistakes in its own report, PIRG surmises that if everyone checked their reports, approximately 133 million reports would contain errors. The reasoning behind this estimate was shown to be unfounded. The study by Arthur Anderson & Co. shows the so-called consumer groups to be grossly mistaken and that serious errors are rare.

The credit reporting industry is a somewhat mysterious entity—one that the average citizen has almost no contact with. This "behind the scenes" aspect makes the industry vulnerable. This paper provides a closer examination of the industry and throws light on the enormous good achieved by credit bureaus, the natural incentives to avoid errors, and the impracticality of expecting infallibility. To

the extent that the critics prevail in influencing laws to hamstring credit bureaus, consumers will benefit very little, but will find credit harder to obtain and available on less attractive terms.

References

- Associated Credit Bureaus, Inc. *Communicator*. (March 1992): 6–7.
 Associated Credit Bureaus, Inc. *Communicator*. (April 1992): 6.
 Arthur Anderson & Co. "Credit Report Reliability Study." 4 February 1992.
 Bibas, Steven A. "A Contractual Approach to Data Privacy." Unpublished Manuscript, Yale Law School, 1992.
 Cole, Ronald H. *Consumer and Commercial Credit Management*. Irwin, Homewood, Ill: 1988.
Consumer Reports. (October 1991): 643–4.
 Consumers Union. "What Are They Saying About Me?" 29 April 1991.
 Dunkelberg, William C., Johnson, Robert W., and DeMagistris, Robin. "Consumer Perceptions of Credit Bureaus." Working Paper no. 26, Credit Research Center, 1979.
 "Food Stamp Waste in '90 Hit \$1 Billion." *Orange County Register*, 15 December 1991, p. A40.
 Elster, Jon. *The Cement of Society: A Study of Social Order*. New York: Cambridge University Press, 1989.
 Hayek, Friedrich A. *Law, Legislation and Liberty*. (3 vols.) Chicago: University of Chicago Press, 1973, 1976, 1979.
 Herubin, Danielle. "Stuck on Report." *Orange County Register*, 2 December 1991, p. C1.
 Klein, Daniel B. "Promise Keeping in the Great Society: A Model of Credit Information Sharing." *Economics and Politics* (1992): 117–36.
 Kurth, Walter R. "On Amendments to the Fair Credit Reporting Act." Testimony before the Subcommittee on Consumer Affairs and Coinage. House. Committee on Banking, Finance and Urban Affairs, 6 June 1991.
 Merry, Sally Engle. "Rethinking Gossip and Scandal." In *Towards a General Theory of Social Control*, pp. 271–302. Edited by D. Black. New York: Academic Press, 1984.
 Mierzwinski, Edmund. "Nightmare on Credit Street or How the Credit Bureau Ruined My Life." Report. United States Public Interest Research Group, 12 June 1990.
 Mierzwinski, Edmund. "Don't Call; Don't Write; We Don't Care: A Survey of Complaints About Credit Bureaus." Report. United States Public Interest Research Group, 6 June 1991.
 Miller, Michael W. "Credit-Report Firms, in Switch, Back New Rules." *Wall Street Journal*, 25 October 1991a, p. B1.
 Miller, Michael W. "TRW to Give Credit Reports Free of Charge." *Wall Street Journal*, 15 October 1991b, p. B1.
 Smith, Adam. *Lectures on Justice, Police, Revenue, and Arms*. 1763. Edited by Edwin Cannan. New York: Augustus M. Kelley, 1964.
 Smith, Adam. *The Wealth of Nations*. 1776. Reprint. Edited by Edwin Cannan. New York: Random House, 1937.

CATO JOURNAL

Smith, Elliot Blair. "TRW to Give Consumers Free Copies of Credit Reports." *Orange County Register*, 15 October 1991, p. B1.

Woodyard, Chris. "TRW Will Offer Consumers Free Report on Credit." *Los Angeles Times*, 15 October 1991, p. D10.

BENEFITS OF MINIMAL LAND-USE REGULATIONS IN DEVELOPING COUNTRIES

David E. Dowall

Introduction

In most developing country cities, the legal, formal sector is largely irrelevant in terms of meeting the basic shelter needs of low- and moderate-income households. Housing that is delivered within the confines of legally sanctioned procedures is normally affordable to those earning at or above the median household income. Thus, those households earning less are forced to look elsewhere for shelter. In virtually all cases, their search leads them to the informal sector, where government rules and regulations associated with formal housing production are, by necessity, ignored. In the most desperate cases, low-income settlers invade land—and no payment is made for plots. While squatting was prevalent in the 1960s and 1970s, it is less widespread today. Nowadays it is more common for informal settlements to take place on illegally subdivided lands which are either rented or sold. The major outcome of these informal approaches is the relatively efficient production of low-cost shelter (Dowall 1991).

For many years, governments and policy analysts viewed these informal settlements as slums needing eradication. Such programs were based on numerous misconceptions about the slums (Hamer 1985). A common belief is that informal settlements are chaotic, posing serious threats to public safety and health. While there are examples of precarious settlements on hillsides and in floodplains,

Cato Journal, Vol. 12, No. 2 (Fall 1992). Copyright © Cato Institute. All rights reserved.

David E. Dowall is Professor of City and Regional Planning at the University of California, Berkeley.

much informal development is planned, following quite acceptable standards.

More recently, policies towards informal housing developments have been shifting. Now it is commonly recognized that informal housing is a valuable capital asset which should not be eradicated (Mayo, Malpezzi, and Gross 1986). Also, and perhaps more importantly, there is a growing recognition that informal-sector housing production is an important overall component of the housing supply system. In the past several years, researchers and policy analysts have stressed the importance of making housing markets work more efficiently by removing burdensome regulations.

The relationship between the informal housing production sector and government regulations over housing and land development is direct and reciprocal. The informal sector exists because of government regulations. Remove them and you will eliminate the blemish of informality. What were one day informal settlements will the next day become low-cost housing subdivisions. Without regulations, the marketplace will determine what households are willing and able to purchase in terms of housing services.

Such an overtly laissez-faire position ignores the fact that housing developments generate significant externalities, such as water pollution, traffic congestion, and soil erosion. Residents of housing projects demand public services, schools, clinics, police protection. The relevant policy question to ask is: what are the minimum levels of regulations or standards which can effectively balance concerns about affordability and access by the poor to housing, with broader community-wide interests? While this short paper cannot possibly provide a definitive answer to this question, it does offer some insights to the costs of high levels of regulations.

How Government Policies and Actions Affect Land and Housing Development

Many cities around the world use master plans, zoning, subdivision regulations, and building codes to control development. These regulations are normally adopted to help protect the urban and natural environment, to gear development with infrastructure capacity, and to maintain and enhance the property values of neighborhoods. Quite often the planning systems and regulations adopted by the local authorities replicate those used in developed countries, with colonial connections usually dictating usage.

In the course of adopting these regulations, little if any thought is given to the potential cost-effects of the controls. For example, how

will minimum lot-size standards affect lot costs? This is unfortunate, since there is ample evidence that excessive land use and development controls drive up housing costs and thereby lower affordability.¹ The remainder of this paper summarizes the available developing countries evidence.

Government regulations over land use and land development affect land and housing markets in three broad ways: (1) land supply constraints, (2) excessive plot-size standards and subdivision design, and (3) procedural delays and red-tape. The following sections take up these three types of constraints.

Land Supply Constraints

Restrictions on the supply of land and the density of residential development greatly affect land costs. As James Ohls, Richard Weisberg, and Michelle White have illustrated, zoning regulations, if they restrict the supply of land available for development below that which would be normally exchanged in the market, operate to increase land prices (1974). The supply of residentially zoned land is often limited when communities attempt to maintain environmental quality or fiscal position by designating land for open space or agricultural use, or for more fiscally desirable commercial or industrial activities. These patterns can be found in developing countries as well, as the following examples illustrate.

Bertrand Renaud (1989) has illustrated the significant impacts of land use controls on land and housing prices in Seoul. Three government policies have constrained the supply of developable land in Seoul: strong zoning policies which restrict the conversion of agricultural land, a greenbelt policy to block the further outward expansion of the city, and land readjustment methods linked with monopolistic administrative practices to force up land prices.

In Seoul, the precipitous increase in land and housing prices challenged the stability of the Roh government. During the late 1980s, land prices in metropolitan Seoul increased at an annual rate of over 25 percent. According to the Korean Research Institute for Human Settlements, the annual increase in land values in 1988 exceeded the annual wage income for all of the country's workers (Clifford 1989).

In Karachi, despite the fact that 90 percent of the land is in public ownership, the supply of *serviced* plots is constrained due to the

¹Considerable research on the effects of land use controls on housing costs has been conducted in North America and Western Europe. See Dowall (1984), Fischel (1990), Fischel (1990), Chesire (1989), and Evans (1988).

lack of infrastructure provision (Dowall 1991). As a consequence, land and housing prices have increased in real terms at 11 percent per year—far faster than household income. This has resulted in the acceleration of informal-sector settlements during the 1970s and 1980s, as more and more low- and moderate-income households were priced out of the housing market. Katchi abadi (informal unplanned) areas accounted for 33 percent of residential land conversion taking place in Karachi between 1970 and 1987.

In India, urban land use controls and policies have a dramatic impact on land supply and price. India has had land use planning controls since the 1950s. In the 1960s, policies were expanded and urban renewal schemes and public development authorities were established. In 1976, the Urban Land Ceiling Act was adopted in an effort to check speculation. As a World Bank report indicated,

Past urban land management strategies have not been overwhelmingly successful in meeting the more important objectives outlined by the 1965 Committee on Urban Land Policy: providing adequate quantities of urban land at reasonable prices and safeguarding the rights of the underprivileged [Wright, Sherer, Hamer, and Bertaud 1984].

One of the most alarming trends is the rapid growth of the slum areas. As of 1983, India's slum population stood at between 32 and 40 million people, and was growing considerably faster than the overall urban population. Land price inflation has been enormous. In Bombay, real land prices increased by 720 percent between 1966 and 1981 (Wright, Sherer, Hamer, and Bertaud 1984). The Urban Land Ceiling Act has caused substantial problems—significant reductions in the supply of land for residential development, creation of a vast black market for real estate, and an overall worsening of housing affordability in India's major urban areas (Acharya 1989).

In Serpong, Indonesia, a suburb southwest of Jakarta, the city's currently adopted land use and zoning plan allocates only 34 percent of the total planned area for residential development. In total, only about 15 percent of the residential area is accessible and has infrastructure. Thus, the actual developable land in Serpong is limited to less than 30 square kilometers. Most of the land is set aside for agricultural and open-space uses, roads, and nonresidential activities (Bertaud 1989). Serpong's plan, like those found in many other cities, ignored reality. While plans are prescriptions of what should or ought to be, government officials frequently treat the plan as given, and program infrastructure into areas where there is limited demand. In Serpong and elsewhere, areas designated as residential development

frequently have no settlements and many areas zoned for agriculture or open space are developed with informal residential subdivisions.

Subdivision Standards

Throughout the Third World, the most common problem is high subdivision standards. In case after case, municipal governments set very high regulations for subdivision layouts. The net result of these high levels is that the minimum costs of plots are normally beyond what households can afford.

In addition to reducing the supply of land for residential development, regulations covering land development standards restrict the intensity of development by requiring large plot sizes or excessive amounts of land for circulation and open space within subdivisions. Large lot-size requirements increase the minimum price of residential lots. While large-lot zoning reduces the per-acre price of raw land, the reduction in prices is often offset by higher land requirements.

In Karachi, excessive land subdivision regulations stipulate large residential plots. All of the plots allocated by the Karachi Development Authority, the city's largest land developer, are over 60 square yards. In Malaysia, land use regulations and standards add considerably to housing costs. The area per house provided for roads in the typical Malaysian subdivision is up to four times greater than in comparable North American or Western European projects. According to accepted international practices, about 25 percent of the land set aside in the average subdivision is wasted. The streets are too wide, the setbacks too great, and land is set aside for redundant community facilities (World Bank 1989).

A recent World Bank report examines the implications of increasing the efficiency of land use and infrastructure standards (Bertaud and Wright 1988). In terms of subdivision standards, there are four factors which influence the cost of plots: (1) plot frontage, (2) block length, (3) street width, and (4) infrastructure standards.

For plot frontage, two competing factors work to determine cost and price: with narrow lots, more lots can be subdivided within each block. This reduces the per-plot costs of streets and infrastructure. On the other hand, narrow plots mean that the houses would have narrow rooms and be less attractive to potential buyers. Thus, prices for very narrow plots are quite low. Based on simulations of the costs to build a hypothetical but realistic housing project with plot size averaging 35 square meters, the most profitable frontage width is between 4.0 and 4.25 meters.

Variations in block length can significantly impact subdivision costs as well. This is because there are economies in road space and

infrastructure deployment as the overall length of blocks is increased. Road width impacts plot costs too: wider roads are more expensive to construct per plot, and wider roads take up more space and make less of a subdivision marketable. Changing standards for infrastructure, such as the types of permitted road surface materials or the diameter of water pipes, can also influence the costs of plot developments.

In the case of Malaysia, excessive subdivision standards pertaining to plot sizes, setbacks, street widths, community facilities, and retention ponds limit the amount of a subdivision's land that can be marketed to between 28 and 47 percent (World Bank 1989). This range is far lower than found in other countries (60 to 70 percent), making housing costs extremely sensitive to land costs.

An assessment of land subdivision regulations in Uttar Pradesh vividly illustrates how lot costs can be dramatically reduced by lowering standards (Bertaud, Bertaud, and Wright 1988). Furthermore, lower standards can legitimize the informal production of plots and make it easier for informal-sector developers to deliver plots in the marketplace.

Besides raising the costs of plot development, land subdivision regulations limit the ability of developers to respond to rising land costs by altering the design of subdivisions. As land prices increase, strict plot-size or circulation requirements make it difficult to build at higher densities.

An example of such flexibility is illustrated in Bangkok, where land prices have increased dramatically over the past three years. The real price of serviced and unserviced residential plots increased by 21 and 37 percent per year respectively between 1988 and 1990. As a result, developers in Bangkok have dramatically shifted their production of affordable housing from townhouses to condominium units (Dowall 1992). Projects are denser and on smaller sites. Developers in Bangkok, like elsewhere, are market-driven—they build housing that is profitable to provide. When unconstrained by regulations, they will respond with product that is attractive to the consumer.

Procedural Delays and Red Tape

The final way in which government regulations influence land and housing costs is through regulatory complexity. Complicated procedures for obtaining development permission make it difficult for developers to respond quickly to changing housing demands and create barriers for new firms wanting to build and sell housing.

A comparison of Malaysia's and Thailand's system of development approval is instructive. A recent appraisal by the World Bank concluded that newly-built housing prices in Malaysia increased by an annual rate of 18.9 percent between 1972 and 1982, a rate about triple the overall increase in consumer prices (World Bank 1989). According to the Bank report, the reason for the rise in Malaysia's housing prices is the combination of high government-imposed housing standards, overly complex and time-consuming housing project approval procedures, the sluggish response of the housing industry to increases in housing prices, and high housing demand. For example, it takes between five and eight years to obtain all the necessary permits from 15 to 20 government agencies for subdivision approval. In Thailand, in sharp contrast, it takes about five months to secure subdivision approval from five government agencies.

With a five- to eight-year lag in housing supply response, housing prices will increase rapidly with a growing demand. The World Bank concluded that Malaysia's most critical housing policy issue is the lack of housing supply responsiveness to price increases.

Steps Toward Market Liberalization

Land use regulations powerfully shape the process and cost of land and housing production. They impede land supply, increase the costs of residential plots, and limit the ability of developers to respond to housing demand. Complicated approval procedures limit housing market competition by creating barriers to entry. They also make the housing market more sluggish and slow to adjust production to meet changes in demand. All in all, the regulations make the threshold price of housing more expensive than it might be otherwise.

The paper raises two fundamental public policy questions. The foremost one is: should governments be legislating standards that less than 50 percent of their citizens can afford? The answer is *no*. The second question, which follows from the first, is: how should governments go about liberalizing land and housing markets? The remainder of the paper turns to this question.

Most governments are "flying blind" when they prepare land use plans and set subdivision controls (Dowall 1990). They simply do not know what is going on in their local land market. Visits to the planning offices of most large third-world cities reveal how little is known about patterns of urban land development, the number of housing units (both formal and informal) built in the past year, land and housing prices, rents for office buildings and factories, infrastruc-

ture deployment patterns, land subdivision patterns, and so on. Given the important role that governments play in shaping land market outcomes, it is extremely important for the implications of their investment and regulatory decisions to be understood.

In order to eliminate the unfortunate impacts of well-intentioned regulations and policies, governments need to undertake a thorough assessment of the urban land markets of their major cities and towns. A land market assessment (LMA) serves as an important first step for reforming government regulations (Dowall 1990). The LMA can be used to answer questions such as:

- Is the supply of urban serviced land expanding to meet growing population and employment needs?
- Which land uses are growing the fastest?
- Where is urban land conversion taking place?
- Where is urban land conversion outstripping the supply of serviced land?
- Are land prices increasing faster than the overall rate of inflation?
- Where are land prices the highest, and where are land prices increasing the fastest?
- How much land is being provided with minimum services needed for future urban development?
- Is enough infrastructure being programed to accommodate urban growth for the next five years?
- Is the price and affordability of housing and commercial and industrial space changing—are real occupancy costs greater now than before?
- Are planning standards and building codes pushing up housing prices?
- Which segments of the population do not have access to housing produced by the formal private sector?
- Are specific public policies or actions constraining the land market?

Land market assessments can also be used to provide baseline estimates of future urban land requirements. They can help guide infrastructure programing and investment decisions and develop land use planning policies. For example, LMAs can be used to estimate the demand for residential plots and commercial and industrial space requirements associated with projections of population and employment. In San Pedro Sula, Honduras, a strategic land development process was designed to promote the supply of land for future growth.

Having the information about the linkages between land supply and subdivision regulations is the first step towards land and housing

market liberalization. The second step is to determine how design and infrastructure standards can be reduced. Here, some technical assistance to both public- and private-sector planners will be needed (see Carroll 1980). Such technical assistance needs to target on senior policymakers and politicians if change is to occur. On standards, the technical assistance needs to span design and site planning, infrastructure engineering, and budgeting.

The levels of standards are not solely technical issues. As Ralph Gakenheimer and Carlos Brando (1987) point out, standard setting is mostly an institutional issue. They suggest four reform areas: (1) reorganize the way we plan, design, build, and operate standards, (2) reorganize the budgeting process, (3) revise the social contract for infrastructure to emphasize concern for welfare and development, and (4) show the savings and improvement resulting from more appropriate standards.

Another set of questions regarding the appropriate level of standards centers on whether design and infrastructure standards should vary across neighborhoods and cities to allow for lower-cost developments, and whether standards can be designed to start low and then increase over time as the income and ability to pay of subdivision residents increases. The relevant question is what level of standards can people afford, not what is the "right" standard.

A third step down the road to reform is to rethink the role of public and private land development, especially in countries where land development is dominated by the public sector (China, Pakistan, CIS, Poland, and Hungary). In study after study, the role of public land development is identified as a critical constraining factor in limiting the responsiveness of the land and housing market to demand. This third step is a big one, testing the abilities of governments to privatise public land development agencies and promote vigorous competition in land and housing markets among private enterprises.

The fourth and final step is what to do about the existing stock of informal housing—how to regularize it in a cost-effective and cost-recoverable manner. Considerable attention has focused on this issue. In the future, more emphasis needs to center on overall land and housing market liberalization.

References

- Acharya, Ballabh Prasad. "Application of Land Management Tools in Combination: Utilizing the Indian Urban Land and Ceiling Act and the Plot Reconstitution Techniques." *Land Development Studies* 6 (1989): 129–46.

- Bertaud, Alain; Bertaud, Marie-Ange; and Wright, James O., Jr. "Efficiency in Land Use and Infrastructure Design: An Application of the Bertaud Model." *Infrastructure and Urban Development Department Report, INU 17*. The World Bank, Washington, D.C., 1988.
- Bertaud, Alain. "The Regulatory Environment of Urban Land in Indonesia: Constraints Imposed on the Poor." Unpublished memorandum, 1989.
- Carroll, Alan. "Pirate Subdivisions and the Market for Residential Lots in Bogota." World Bank Staff Working Paper no. 435, Washington, D.C., 1980.
- Cheshire, Paul, and Sheppard, Stephen. "British Planning Policy and Access to Housing: Some Empirical Estimates." *Urban Studies* 26 (1989): 469-85.
- Clifford, Mark. "Through the Roof." *Far Eastern Economic Review*, 8 June 1989, p. 102-3.
- Clifford, Mark. "Landed in Trouble." *Far Eastern Economic Review*, 14 September 1989, p. 72.
- Dowall, David E. *The Suburban Squeeze*. Berkeley: University of California Press, 1984.
- Dowall, David E. *The Land Market Assessment: A New Tool for Urban Management*. The Urban Management Programme, UNCHS and The World Bank, Nairobi and Washington, D.C., 1990.
- Dowall, David E. "Comparing Karachi's Informal and Formal Housing Delivery Systems." *Cities* 8 (3) (1991): 217-27.
- Dowall, David E. "A Second Look at the Bangkok Land and Housing Market." *Urban Studies* 29 (11) (1992): 25-37.
- Evans, Alan. *No Room! No Room! The Costs of the British Town and Country Planning System*. Institute of Economic Affairs, London, 1988.
- Fischel, William A. "Do Growth Controls Matter? A Review of Empirical Evidence on the Effectiveness and Efficiency of Local Government Land Use Regulation." The Lincoln Institute of Land Policy, Cambridge, Mass., 1990.
- Gakenheimer, Ralph, and Brando, Carlos. "Infrastructure Standards." In *Shelter, Settlement & Development*. Edited by Lloyd Rodwin. London: Allen & Unwin, 1987.
- Hamer, Andrew Marshall. "Bogota's Unregulated Subdivisions: The Myths and Realities of Incremental Housing Construction." World Bank Staff Working Paper no. 734, Washington, D.C., 1985.
- Mayo, Stephen; Malpezzi, Stephen; and Gross, David J. "Shelter Strategies for the Urban Poor in Developing Countries." *World Bank Research Observer* 1 (2) (1986): 183-203.
- Ohls, James; Weisberg, Richard; and White, Michelle. "The Effects of Zoning on Land Value." *Journal of Urban Economics* 1 (1974): 428-44.
- Renaud, Bertand. "Compounding Financial Repression With Rigid Urban Regulations: Lessons of the Korean Housing Market." *Review of Urban and Regional Development Studies* 1 (1989): 3-22.
- World Bank. *Malaysia: The Housing Sector, Getting the Incentives Right*. Infrastructure Division Country Department II, Asia Regional Office, Washington, D.C., 1989.

MINIMAL LAND-USE REGULATIONS

Wright, James; Sherer, Samuel; Hamer, Andrew; and Bertaud, Andrew.
“India Urban Land Management Study.” World Bank internal memorandum, Washington, D.C., 1984.

TRADE BARRIERS IN THE THEORY OF INSTRUMENT CHOICE

Jean-Luc Migué

The process of choosing among policy instruments has been analyzed mostly as a problem of public choice (Buchanan and Tullock 1975, Cassing and Hillman 1985, Dewees 1983, Borcharding 1983, and Becker 1983) or as a problem in the economics of political institutions (Campos 1989). An alternative approach, not inconsistent with either framework and in fact preceding both, would consist in analyzing the economic or cost constraints facing decisionmakers as they undertake the task of selecting instruments. A first step in assessing the impact of constraints on policy choices has been offered recently by Robert Hahn (1990). Our purpose in the subsequent analysis is to build on his work and offer a formal framework for analyzing the effect on domestic policy choices of constraining the power of national governments to maintain trade barriers.

Domestic Instruments as Complementary to Trade Barriers

A general conclusion derived from Hahn's latest approach has been that the various instruments available to national authorities are substitutes. The logic behind this conclusion is as follows. Suppose that there are two instruments available to legislators for pursuing their objectives as shown in Figure 1: spending, *S*, and regulation, *R*. We assume that these outputs consist mainly of transfers to rival groups for the purpose of maximizing votes. By analogy with the traditional analysis of consumers' or producers' behavior, the combination of policies available to politicians has been represented by the straight budget line, *AB*, in Figure 1. Given the assumption that

Cato Journal, Vol. 12, No. 2 (Fall 1992). Copyright © Cato Institute. All rights reserved.

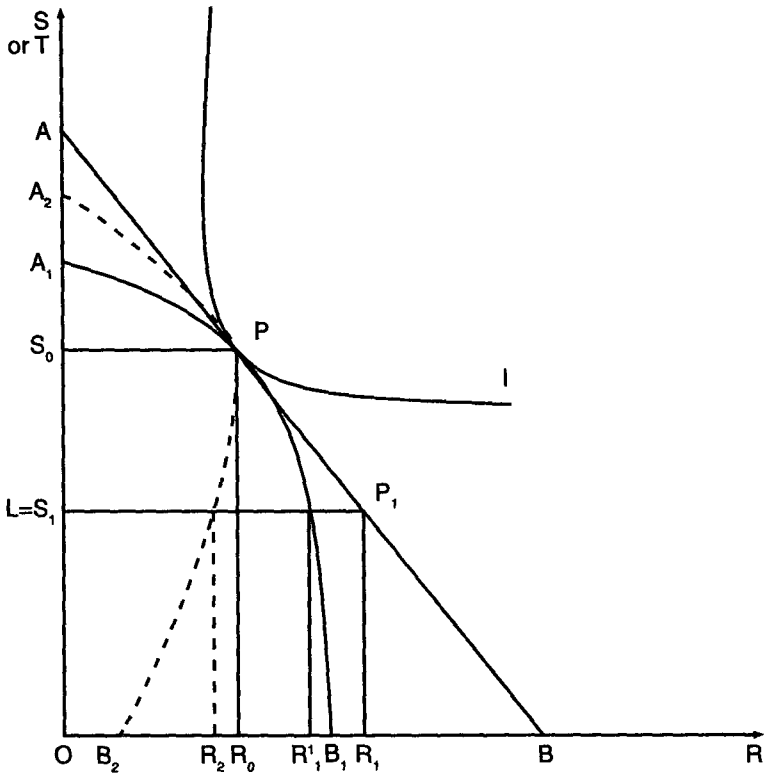
The author is Professor of Economics at the École nationale d'administration publique in Quebec City, Canada.

legislators are only constrained by the transformation curve between the two instruments, they choose point P on their political budget line. Tangency to the indifference curve I indicates that maximum political benefits are obtained from using a spending level of OS_0 in combination with a level of regulation equal to OR_0 . In this Hahnian version of the model, imposing further constraints on policy choices, say a spending limit set at $L=S_l$, causes an increase to R_l in the level of regulation. In more concrete terms, as budgetary limits and treaty provisions prevent politicians from further indulging in farmers' subsidization, supply management of farm production can be reverted to. Coercive provision of day-care services by employers can be substituted for family allowances and provision by monopoly state suppliers. Mandatory health services imposed on employers can be substituted for public medicare with similar redistributive results.

To bring out the implications on national governments of an exogenously imposed removal of trade barriers, two changes must be made in this presentation. The first one is a correction in the implicit assumption made concerning the shape of the political budget line. Realistically this line can only be concave to the origin as represented by the curve A_lB_l in Figure 1. The reason for this change is straightforward. The government has a monopoly on the use of policy instruments. It is not simply a marginal user, as is the case when the graphical presentation is applied to the analysis of individual producers facing the task of choosing between two inputs. Therefore, as greater use of an instrument is made relative to the others, its productivity is bound to decline. Its marginal contribution to the political output or to whatever redistribution is sought by legislators can only decrease. While this modification is unlikely to be strong enough to reverse the direction of the change in the use of other instruments, it is seen at point R'_l that the increase in the use of regulation as a result of constraints on spending is less pronounced than predicted by Hahn. In general, constraints on one instrument result in less significant increases than assumed in the use of others.

Our second modification is more fundamental and bears on the specific properties of one instrument, namely protectionism. The main contention of this paper is that in the arsenal of redistributive instruments available to public authorities protectionism stands out as fundamentally different from all others. I argue that domestic policy instruments and trade barriers are complements rather than substitutes. More specifically, I show that as a tool to redistribute wealth between various groups trade barriers are a prerequisite to the use of all redistributive tools.

FIGURE 1
TRADE-OFF BETWEEN INSTRUMENTS



As exogenous constraints are applied to the use of tariff or nontariff barriers, the response of local and national legislators has to be to reduce rather than increase resort to other instruments. In terms of Figure 1, the political budget line should be drawn as A_2B_2 . As national governments are forced to abandon more and more of their trade restrictions, the budget constraints shift inward and the trade-off line with other instruments becomes upward-sloping. Without the power to close their borders to foreign competition, national governments should find out that most other tools formally available to them to redistribute wealth are of little value. Assuming that trade barriers, T , rather than spending levels are represented on the y axis, it is seen that the level of regulation declines from R_0 to R_2 , when limits $L=S_1$ are placed on the power of national governments to resort to protectionism.

This strong result clearly changes the consequences on government size and welfare of imposing constraints on this specific class of instruments. In contrast to the effects predicted in Hahn (1990), constraints on trade barriers unequivocally result in a decline in government size, even when measured by expenditures implied by both the spending and regulation levels. Similarly assuming that government activities consist solely in redistributing wealth through the use of various transfer instruments, economic welfare rises and the size of government declines as constraints are placed on the level of trade barriers. The logic behind this far-reaching conclusion is the simple logic of federalism at the world scale.

Competitive Federalism at the World Scale

The basic principle underlying federalism within traditional federal states or within common market structures is that the burden of policies cannot be transferred to outside producers and consumers. The reason for this is that prices are determined outside these economies. Neither the states in federal structures nor the national governments of common-market members are in a position to influence prices outside their borders.

The above statement is valid even under a protectionist regime. However, one basic outcome differs after the advent of the more open arrangement associated with freer trade. Because of the free movement of resources, any inefficient measure adopted by a national government imposes a heavier burden on the residents of the jurisdiction in question. Opening the economy to freer trade increases the elasticity of supply and demand. Put differently, domestic interventions exert their redistributive influence through changes they bring about in the relative price of goods, capital, and labor within the national territory. Since national governments cannot influence common market prices, the only way for national legislators of having some lasting impact on prices internally is to maintain exchange barriers with the outside world.

Imposing significantly heavier taxes and stricter regulations on productive national resources increases relative production costs and, under free trade, first encourages a more rapid and more pronounced substitution of imported goods and services for local production. Furthermore, insofar as the burden of these measures is more keenly felt by capital and labor suppliers, victims of abuse can move their capital to neighboring jurisdictions and at the limit "vote with their feet" and leave the territory. Because inefficient decisions cost more in a more open economy, the predicted outcome is less interventionism.

Once resources are free to move across national boundaries, the pressure intensifies on national governments to pursue policies that attract people and firms. The list of effects that differentiate situations with free trade from conditions without could be lengthened at will. The message remains the same: Inefficient decisions cost more in a more open economy because goods, capital, and people can more freely move away from the country that imposes unfavorable legislation. National governments have been able to exercise their cartelizing influence over their economies mainly because they have had the power to impede the free movement of resources between their territories and the rest of the world. And as public decisions eventually are reflected in increased taxes or regulations, the overall impact of free trade on national governments is toward reduced and more neutral taxes and regulations than would otherwise be the case.¹

Resort to conventional graphical presentation of protectionist measures in Figure 2 should make our conclusion obvious. Given the conditions in the national economy as expressed by the demand (D) and supply (S) schedules in Figure 2, the adoption of a tax or regulatory measure shifts the domestic supply leftward from S to S_1 , as production costs are increased. With no national tariffs, the world price to domestic purchasers remains unchanged at Op_0 , and domestic production sharply declines from Oq_0 to Oq_1 . On the other hand, should a tariff equal to p_0t_0 be applied on foreign supply, the same domestic cost-increasing tax or regulatory measure could be accompanied by no changes in, or even a higher level of domestic production, as is the case at Oq_2 in Figure 2. In other words, world supply is more elastic than domestic supply.

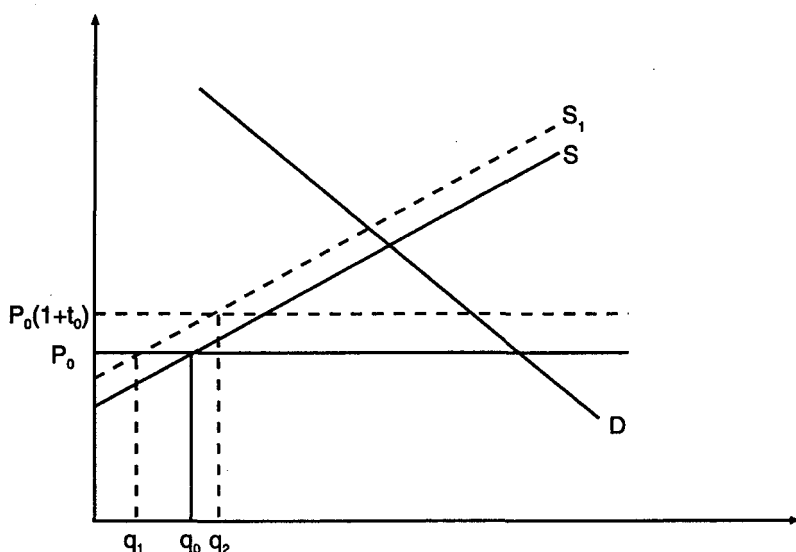
Because they can move their resources away from unfavorable legislation under free trade, resource owners become more sensitive to relative tax and regulatory actions by national governments. The contribution of free trade to solving the problem of the monopolistic state is to limit the power of the government. To analysts of antitrust policy, the analogue with the market power of business firms should be obvious. Just as trade can perform the function of antitrust policy in constraining market power, so can it act as a curb on government power.

Resource Mobility and Taxes as Fees for Services

The single most convincing piece of evidence on the strength of resource mobility is the consistent convergence of per capita income

¹As a further illustration of this process as applied to Canadian conditions, see Migué (1988).

FIGURE 2
IMPACT OF INTERNAL POLICIES WITH AND WITHOUT
FREE TRADE



across regions of the United States and across a broad sample of countries over the past decades. The mobility of goods, capital, and labor has been causally related to this factor (Barro and Sala-i-Martin 1991; Baumol, et al. 1989; and Helliwell and Chung 1991). Convergence occurred in the past despite the numerous hurdles set up by governments against the movement of goods and factors between national communities. The position taken in this paper is that as barriers to trade are removed, tendencies toward income homogeneity between regions are reinforced.

The process of income convergence is an important one. It shows that resource owners in all regions respond to incentives by moving on the basis of compensations received, including regional taxes, regulations, and services. Individuals have been directly shown to be sensitive to the combination of taxes and services offered at the regional level (Cebula 1979, and Cebula and Kafoglis 1986). The rising tax burden is seen by Richard Vedder (1990) as having increased taxpayers' sensitivity to local variations in tax loads. More reliance on the pricing of public services in recent years is interpreted as a consequence of this higher awareness by taxpayers.

After three decades of debate regarding the question of capital mobility as a source of constraints on government arbitrary action, it is generally agreed that tax rate differentials is a determining factor of localization for business firms, within the United States and between the United States and other nations (Benson and Johnson 1986; Hines and Rice 1990; and Harris, et al. 1991). One important mechanism used by firms to move their capital away from jurisdictions with heavier tax burdens is the practice of transfer pricing. It has to do with the price "foreign" companies charge or pay for parts shipped to or from their own "local" affiliates. In order to reduce his reported earnings from plants located in a high-tax country, a foreign owner can overcharge for parts shipped to this local subsidiary. Viewed by local tax collectors as a threat to their tax base, such tax evasion can only work because the foreign owner is based in a low-tax country. In a more analytical perspective, the practice is an important competitive weapon in the arsenal of investors to rein in the tax appetite of national governments in a freely trading world.

Another application of the role of capital mobility in the selection of tax instruments by governments concerns the question of local property taxes as determinants of capital migration. An animated debate has been going on for some time among tax analysts, which is directly relevant to our analysis. Two authorities (Mieszkowski and Zodrow 1983) on the question had previously concluded that real property taxes are distortionary taxes on capital, rather than benefit taxes related to individual consumption of local public goods such as education. Because local authorities find themselves unable to ensure that the last dollar paid in taxes equals the taxpayer's marginal willingness to pay for local services, wholesale free-riding is said to take place on local public goods. Large families can rent lots on the fringe of cities, park trailers on them and send their children to the local school at little cost to them. This result implies that owners of real property in the same neighborhood are charged tax prices higher than the value of local benefits received, which tends to drive capital away from the jurisdiction. Undesirable distortions in the spatial allocation of capital are seen arising from this process.

In the most recent study of this question, William Fischel (1992) has shown that local zoning laws are effective in transforming property taxes into fees for local public services. Municipalities use quantifiable rules such as minimum lot size to prevent unwelcome developers and to minimize rent seeking. Reduced taxation of farmlands and reductions for elderly taxpayers are other examples of adjustments conceded by local officials to residents who do not impose

heavy fiscal burden. Property taxation endures because fiscal zoning transforms it into a benefit tax, protected from wholesale free riding and rent seeking by residents.

Why do towns and cities strive to adjust dollars paid in taxes to the level of benefits received in local services and, in so doing, succeed in overcoming rent seeking by migration? The answer is that they have little latitude to do otherwise. By virtue of the high mobility of resources between their territory and the outside world, tax and regulatory burdens not compensated by benefits in services rapidly develop into movements away from jurisdictions imposing them. Alternatively, should benefits exceed local taxes and regulatory costs, the likely inflow of people or capital attracted by potential gains would tend to dissipate the rent, which is also an unattractive prospect for current residents. In other words, distortionary taxes and regulations are extremely costly to local residents because of the high elasticity of supply of resources. Inefficient migrations into and out of local territories are unlikely to occur.

The ability of decentralized authorities to minimize rent seeking by migrations is in no way limited to municipal jurisdictions. The potential contribution from this alternative solution is evident from real-world illustrations at state and even national levels. With a view to making only its citizens share in the oil rent, the state of Alaska had originally limited the distribution of its tax take on oil to residents who had been in the state before the enactment of the measure. The rejection of the scheme by the U.S. Supreme Court vividly illustrates the power of this process. On the eve of China's takeover of Hong-kong, Canada and other countries offer wealthy Hong Kong residents the opportunity to purchase the right of citizenship by investing a minimum amount of their money into the host economy. Gary Becker (1992) has proposed to institutionalize the immigration process by substituting an auction system for the arbitrary criteria now in place to select immigrants. Because of the lower mobility of its French-speaking population, the Canadian Province of Quebec has had to rely more than other provinces on payroll taxes than on profit taxes to finance its programs.

The end-result of this logic is that under free trade national and local taxes become merely fees for national and local public services. In a world where resource movement is unimpeded by trade barriers, taxes and regulations that do not meet this requirement become so costly to raise that they tend to be abandoned by governments. This conclusion has been forcefully argued by Fischel (1989) in his penetrating interpretation of the famous 1978 Proposition 13 in California. By mandating the strict equalization of educational spending among

California districts, a California court had made it impossible for higher local taxes to buy better schools. The decree converted the local property tax from a benefit tax into a costly distortionary tax on capital. Voters in high-demand school districts chose to reduce drastically their reliance on the property tax in consequence of this trend. By a similar process, the decline of the property tax as a source of local school financing in most Canadian provinces coincided with the rise of the spending equalization movement.

General Implications of World Federalism

As a competitive arrangement rooted in the Tiebout (1956) tradition, freer trade places national governments in the approximate position of a province or a state vis-à-vis the national economy in a federal state. The logic of federalism operates in any political structure where the power of public authorities extends to less than the size of the economy in which resource movement is unimpeded by trade barriers. Its competitive action can be at work within national economies in decentralized federal states or between countries associated in common-market arrangements with only limited central powers. Whether the central layer of government is made up of elected officials as in federal states and in the European community or of multinational bodies as is the case in the Canada-U.S. free-trade area is analytically irrelevant. The determining characteristic of a federalist structure is that most responsibilities are entrusted to decentralized authorities who have no power to tax and regulate the whole area where trade is free. The provision of remaining public goods for the overall community is carried out through a central body of institutions, whose decisions are binding on national authorities and private agents across the whole bloc. International agreements can be viewed as arrangements for the provision of public goods common to more than one country. NORAD, NATO, the OECD, the agreement on acid rain between the United States and Canada, the IMF, even the United Nations or the binational commissions to settle disputes between New Zealand and Australia or between Canada and the United States are no less central entities than the federal government of Switzerland or the Brussels authority. Multinational bodies may even offer the advantage of being less susceptible to the centralist tendencies of formally federal structures.

The market, in general, acts as a mechanism to reveal preferences through the mobility it gives players to enter and exit. Politics relies on "voice" to discipline governments, but without mobility (Hirschman 1970). Restricting citizens' movements is an instrument of state

monopoly. Opening the way for the mobility of goods, capital, and persons between regional and national communities, serves as a substitute to politics for expressing preferences where public goods are concerned. This federalist solution is the counterpart of the market in the public sector. It is the institutionalizing of competition between governments. It offers individuals an additional instrument to discipline governments. In a real sense, free trade at the common-market level is a substitute to national bills of rights as a guarantee of individual freedom and prosperity. Freer trade brings not only prosperity to the interacting economies, it offers more freedom and the promise of a more efficient provision of public services. The essential condition for this desirable result is that the flow of trade, capital, and people remains unimpeded between national boundaries.

Centralization versus Federalism

In a decentralized, competitive system there is no need to limit legally or constitutionally the legislative power of the various national governments, provided the mobility of resources is not impeded. Exit not voice is the instrument used by individuals to make their preferences known. But by the same logic, exit will only be an effective means if national decisions are not superseded by a vast central power covering the same fields, within trade blocs or at the supra-national level. Centralization is defined as the power of a central authority to rule over the entire common-market economy. As such, it weakens the power citizens have over their government and opens the way for interest groups and rent seekers to obtain transfers at the expense of the population as a whole. What use is the freedom to leave a country, what use is the choice that consumers have to purchase their goods from outside suppliers, if the central government can regulate and tax the economy at the common-market level? Individuals and groups bound by one region no longer have the option to escape its restrictions by emigrating to another region, as the federalist solution is thwarted by the cartel implemented at the central level. It is no accident that throughout history dictatorial regimes have deemed it necessary to tightly control the movement of resources across their borders.

Paradoxically, in the conventional debate in the various federal states and in Europe, federalism is associated with the strengthening of the central government's powers, while the search for autonomy and decentralization is linked to narrow secessionist movements and anti-European options. Supporters of devolution are immediately

labelled anti-Europe and “nation-builders.” Yet the contrary is true. Federalism is decentralization. As a form of government, federalism is the opposite not of decentralization but of the unitary state. The partisans of devolution are the real federalists. The assessment of a conventional analyst (Eichengreen 1991) on the Savings and Loans crisis in the United States illustrates the confusion. In a perfectly perverse use of language, the author designates by the label “fiscal federalism” the arrangement whereby depositors in Texas received \$20 billion in deposit insurance, while the U.S. Treasury collected only \$1.3 billion in taxes from S & Ls in that state. Two implicit assumptions are contained in this judgment. First, federalism is measured not by the extent of decentralization and competition but by its opposite, the power of a central authority to neutralize federalist forces. Second, it is wrongly assumed that the government of a unitary state could not have done even better in terms of redistribution in favor of some regional interests.

It is regrettable that this centralist vision has become mainstream teaching among economists. In their assessment of the proposed unified European currency, two authorities (Sachs and Sala-i-Martin 1992) argue that such a plan could result in very serious strains unless some type of EC-wide tax-transfer scheme is put in place. Their idealized model is the United States of America where the federal government absorbs, in reduced taxes and increased transfers, up to 40 percent of any decline in a typical region’s income. In contrast, the European Community, they lament, merely cushions 0.5 percent of the impact of an adverse shock. Nowhere in this evaluation is it recognized that a currency union combined with subsidies does not suppress the cost of regional or national inconsistent policies. It merely shifts the cost of adjustment to the more conservative and prosperous regions of the union.

Balkanizing Common Markets by Centralization

An idea widely held by politicians, bureaucrats, and some economists is that only a strong central authority can safeguard the economic union of a common market, such as Europe or North America, against the narrow protectionist forces of local and national interests. History and analysis teach a totally different lesson. They show that the central government of existing federations or a strong central authority in Europe could and do effectively use their greater monopoly power over the common market to balkanize the overall economy in erecting insuperable inter-regional barriers between member states. If one defines balkanization as a process whereby a govern-

ment intervention dissociates the price or tax burden of regional goods and services from their production cost, then balkanizing a common market is equivalent to erecting tariffs, quotas, and protectionist subsidies between the member states of a common market (Migué 1993).

Central supply by a supra-national authority can be substituted for national supply, directly or indirectly, as when equalization handouts are offered and costs are shared with member states. The allocation method favored by central governments is service uniformity over the whole territory at a quality level approximating the preferences of the median voter for the community as a whole. The centralized provision of public services then acts as implicit subsidies to less developed members of the common market. Selling public services on the cheap to lower-income countries is no different from offering discriminatory subsidies to national producers or exporters in those countries. In reducing regional production costs, the practice hinders the process of resource specialization and acts as a protectionist measure in favor of lagging areas. By contrast, production costs are raised in the most productive countries as a result of the increased tax burden.

In addition to setting up a central monopoly supplier, two further instruments are available to the supra-national government of a common market to pursue its discriminatory pricing goals: (1) payments to decentralized governments, and (2) central regulation. The uniform tax regime implemented at the common-market level by equalization payments or cost-sharing arrangements, means that productive resources in prosperous areas have no incentives to move away from jurisdictions burdened by their excess share of the common programs. Alternatively, producers in subsidized territories are encouraged to stay in their less productive employment by what has become in effect a cartel of member states at the common-market scale. Central handouts cause national or regional prices to be dissociated from regional costs. They inhibit the specialization of the community's resources and are protectionist in nature.

The irony of the proposed centralist vision of Europe is that while the central provision of public services and the contemplated cost-sharing and regional-development policies would result in lower resource mobility and excess population in declining areas, social and economic regulations by the central authority would imply reduced investment and population in less prosperous countries. The proposed social charter with its alleged rights to standardized health, education, and social-insurance services, with its support for uniform environmental standards, with its goal of imposing common

union wages and working conditions (affirmative action programs), would have two important consequences. First, it would cause production costs to rise throughout the territory. Second, the common standards would clearly be set by the demand of interest groups in the most developed and prosperous urbanized regions. The ability of lower-income, lagging areas to compete with the most productive ones would be reduced in the same proportions. Poorer regions would be denied the power to compete with prosperous ones through lower wages, lower taxes, or fewer environmental amenities.

Redistributive policies of central governments foster resistance to necessary regional adjustments in another way. By releasing member states from the consequences of their decisions, central subsidies and regulations encourage provincial or national governments to show little concern for adopting adjustment policies. Equalization payments, cost-sharing programs, and central regulatory measures mostly serve to shield decentralized administrations from the consequences of their own cost-increasing policies. They act as cartel enforcers for regional governments. All three instruments result in policy costs being shifted to other parts of the common market. The more inefficient national or regional governments are, the more they are compensated by the central authority. This is nothing less than an exercise in economic absurdity.

The conclusion to be drawn from past and planned practices of centralized structures is unambiguous. If the distortion of prices in member states is accepted as a valid measure of protectionism, then federal initiatives which impose the uniform treatment of regions in matters of taxation, expenditures, and regulations are distortionary and protectionist in nature. All central interventions with variable local incidence are the equivalent of implicit regional customs tariffs, regional quotas or discriminatory subsidies to regional consumers and producers. They all inhibit the specialization of the community's resources and restrict trade. To argue against decentralization on the basis of the requirements of economic integration within common markets amounts to an erroneous evaluation of the facts.

Evidence

Ironically little effort has been devoted so far to assessing methodically the relationship between protectionism and domestic interventionism. While numerous studies have been undertaken to analyze the determinants of trade barriers on the one hand and the size of the public economy on the other, few analysts have attempted to relate the two phenomena. One study that comes closest to tackling

the kind of questions that concern us here is an empirical study by John Conybeare (1983). He finds that tariff rates are positively correlated across countries with the size of the central government measured as a percentage of total government. This striking result is all the more significant since David Cameron (1978) has shown that the expansion of the public economy across countries is largest in unitary, highly centralized nations. Decentralization at the national level tends to dampen the degree of expansion of the public sector.

Broad historical trends support the hypothesis that the share of the government sector in the national economy is positively associated with the levels of trade barriers. Until the 1930s (mostly under the gold standard), there were almost no restrictions on the movement of people and capital in the industrial world. Money and goods moved in large relative volume. It has taken some 70 years for merchandise trade as a proportion of GNP to overtake levels it had achieved prior to World War I, with the bulk of the movement occurring between 1950 and 1970. Few quantitative barriers to trade were raised and few obstacles impeded capital flows. Nontariff barriers are of comparatively recent popularity and were practically unknown. Countries in Western industrial economies had low or nonexistent tariffs, including Great Britain. On the other hand, data now available on the long-term trend in government expenditures in the United States, the United Kingdom, Denmark, and Sweden, show that the size of the central government remained practically constant at 10–15 percent of GNP over the whole industrial era prior to the 1930s, and then grew sharply and consistently until very recent periods (Tullock 1990). In this paper Tullock argues that the broad association between trends in levels of protectionism and government expenditures is not coincidental: The turning point coincided with the onslaught of protectionism in the 1930s. This result is consistent with the thesis that the spending instrument (and regulatory tools) can only be used on a large scale when trade barriers are steep.

It is interesting to note that some writers are detecting a reversal of the trend in the growth of government expenditures and taxes in major industrial countries (McKenzie 1988, and McKenzie and Lee 1991). Outlays seem to have peaked in the early 1980s, sometimes earlier. Tax rates are seen declining in most industrialized countries. More important, many governments followed the United States example in lowering their highest marginal tax rates on income. Two other taxation trends, well documented in Canada (Bélanger 1991) and presumably observed elsewhere, underlie the competitive forces at work in more open economies in the last 30 years, namely the rise in the taxation of wages relative to capital, and the lightening

of the tax burden on savings relative to consumption. Dampening tax-induced disincentives to save and to grow is what one would expect from lower trade barriers and the incident competition from abroad. The worldwide movement toward deregulation and privatization in the late 1970s and 1980s can be viewed as consistent with the greater mobility of resources coincident with reduced trade barriers.

What is remarkable in these statistics from our standpoint is that the start of the new trend follows five rounds of successful multilateral negotiations under GATT and the advent of the European Community. In the meantime, Canada and the United States have negotiated a free-trade agreement, as have the members of the Mercosur group (Argentina, Brazil, Paraguay and Uruguay). Similar trends in other regional blocs are observed as protectionism gives way to more formal commitments to lower barriers within a Japanese-led group in Asia and Mexican-led groups in Central and North America. The five countries of the 1960 Andean Pact (Peru, Bolivia, Ecuador, Colombia, and Venezuela) have also relaunched their plans for economic integration. In the same vein Eastern Europe can be viewed as moving toward a free enterprise zone. Instead of a small part of a country being freed, entire countries, at one fell swoop, are being freed from constraints, while little central power appears to be in demand.

These are clearly rough estimates of the role of trade barriers on national public economies. Nontariff barriers are not incorporated into the picture, and the extent of interventionism is limited to budget dimensions, with the whole package of national regulations left out. More empirical work remains to be done before the picture gets clear. It is reassuring at this stage that casual empiricism of the kind used above is not inconsistent with the theoretical predictions, but on the whole supportive.

Conclusion

This paper has proposed a formal framework for analyzing the effect on domestic policy choices of constraining the power of national governments to maintain trade barriers, as experienced in GATT-type arrangements, in common-market treaties, and in other free-trade agreements within blocs of trading partners. Without the power to close their borders to goods and services, to capital, and ideally to people, national and regional governments should find out that resort to domestic policy instruments, such as spending, taxes, and regulations, for redistributive purposes is more costly, indeed

impossible. A smaller share of resources will be transferred to governments under free trade.

References

- Barro, Robert J., and Sala-i-Martin, Xavier. "Convergence across States and Regions." *Brookings Papers on Economic Activity* 1 (1991): 107–82.
- Baumol, William J., et al. *Productivity and American Leadership: The Long View*. Cambridge, Mass.: MIT Press, 1989.
- Becker, Gary S. "A Theory of Competition among Pressure Groups for Political Influence." *Quarterly Journal of Economics* 98 (August 1983): 371–400.
- Becker, Gary S. "An Open Door for Immigrants—The Auction." *Wall Street Journal*, 14 October 1992, p. A14.
- Bélanger, Gérard. "Évolution de la taxation au Canada." *L'Analyste* 35 (Fall 1991): 47–49.
- Benson, Bruce L., and Johnson, Ronald N. "The Lagged Impact of State and Local Taxes on Economic Activity and Political Behavior." *Economic Inquiry* 24 (July 1986): 389–401.
- Borcherding, Thomas E. "Toward a Positive Theory of Public Sector Supply Arrangements." In *Crown Corporations in Canada*, pp. 99–184. Edited by Robert S. Prichard. Toronto: Butterworth & Co. (Canada) Ltd, 1983.
- Buchanan, James M., and Tullock, Gordon. "Polluters' Profits and Political Response: Direct Controls Versus Taxes." *American Economic Review* 65 (March 1975): 139–47.
- Cameron, David R. "The Expansion of the Public Economy: A Comparative Analysis." *American Political Science Review* 72 (December 1978): 1243–61.
- Campos, Jose Edgardo L. "Legislative Institutions, Lobbying, and the Endogenous Choice of Regulatory Instruments: A Political Economy Approach to Instrument Choice." *Journal of Law, Economics, and Organization* 5 (Fall 1989): 333–54.
- Cassing James H., and Hillman, Arye L. "Political Influence Motives and the Choice Between Tariff and Quotas." *Journal of International Economics* 19 (November 1985): 279–90.
- Cebula, Richard J. *The Determinants of Human Migrations*. Lexington, Mass.: Lexington Books, 1979.
- Cebula, Richard J., and Kafoglis, Milton Z. "A Note on the Tiebout-Tullock Hypothesis: The Period 1975–80." *Public Choice* 48 (1986): 65–69.
- Conybeare, John A.C. "Tariff Protection in Developed and Developing Countries: A Cross-sectional and Longitudinal Analysis." *International Organization* 37 (Summer 1983): 441–67.
- Deweese, Donald N. "Instrument Choice in Environmental Policy." *Economic Inquiry* 21 (January 1983): 53–72.
- Eichengreen, Barry. "Is Europe an Optimum Currency Area?" NBER Working Paper no. 3579, January 1991.
- Fischel, William A. "Property Taxation and the Tiebout Model: Evidence for the Benefit View from Zoning and Voting." *Journal of Economic Literature* 30 (March 1992): 171–77.

- Fischel, William A. "Did Serrano Cause Proposition 13?" *National Tax Journal* 42 (December 1989): 465-75.
- Hahn, Robert W. "Instrument Choice, Political Reform and Economic Welfare." *Public Choice* 67 (December 1990): 243-56.
- Harris Davis, et al. "Income Shifting in U.S. Multinational Corporations." NBER Working Paper no. 3924, December 1991.
- Helliwell, John F., and Chung, Alan. "Are Bigger Countries Better Off?" In *Economic Dimensions of Constitutional Change*, pp. 345-67. Edited by Robin W. Boadway, Thomas J. Courchene, and Douglas D. Purvis. Kingston, Ontario: John Deutsch Institute for the Study of Economic Policy, 1991.
- Hines, J. R., and Rice, Edward M. "Fiscal Paradise: Foreign Tax Havens and American Business." NBER Working Paper no. 3477, October 1990.
- Hirschman, Albert O. *Exit, Voice and Loyalty*. Cambridge, Mass.: Harvard University Press, 1970.
- McKenzie, Richard B. "The Twilight of Government Growth in a Competitive World Economy." Cato Institute Policy Analysis no. 111, 19 August 1988.
- McKenzie, Richard, and Lee, Dwight R. *Quicksilver Capital*. New York: The Free Press, 1991.
- Mieszkowski, Peter, and Zodrow, George R. "The Incidence of the Property Tax: The Benefit View Versus the New View." In *Local Provision of Public Services*, pp. 102-29. Edited by G.R. Zodrow. New York: Academic Press, 1983.
- Migué, Jean-Luc. *The Balkanization of the Canadian Economy*. Vancouver: The Fraser Institute, 1993 (forthcoming).
- Migué, Jean-Luc. "True Independence Through Free Trade." In *Trade-offs on Free Trade*, pp. 446-53. Edited by Marc Gold and David Leyton-Brown. Toronto: Carswell, 1988.
- Sachs, Jeffrey, and Sala-i-Martin, Xavier. "Fiscal Federalism and Optimum Currency Area: Evidence for Europe from the United States." NBER Working Paper no. 3855, 1992.
- Tiebout, Charles M. "A Pure Theory of Local Expenditures." *Journal of Political Economy* 64 (October 1956): 416-24.
- Tullock, Gordon. *Government Growth*. Mimeograph, University of Arizona, March 1990.
- Vedder, Richard. "Tiebout, Taxes and Economic Growth." *Cato Journal* 10 (Spring/Summer 1990): 91-10.

GOVERNMENT AND RECYCLING: ARE WE PROMOTING WASTE?

Clark Wiseman

Introduction

The municipal solid waste disposal issue has come to be widely regarded as one of the larger ecological and economic problems of the present day. It is accordingly grouped with a variety of similar “crises” that call for substantial business and household behavioral adjustments, reinforced, in cases where the requisite behaviors are not forthcoming, by vigorous governmental programs and incentives. Here, the contrary position will be taken that the municipal solid waste disposal issue is political in nature—a manifestation of a flawed local government decisionmaking process. The point will be further pressed that, far from being a solution to the solid waste problem, the nation’s massive recycling effort—sustained and expanded by subsidies, taxes, and government operation—is itself inherently wasteful.

At the outset, it is emphasized first that this paper deals only with municipal solid waste, to the exclusion of industrial and hazardous waste. The discussion is focused upon the issues raised by the gross volume of nonhazardous waste generated and disposed of or recovered, rather than the important but distinct set of issues relating to recycling or disposal of specific toxic sources such as batteries, motor oil, paints, and household chemicals. Second, at present, hundreds of proposals are pending before the state legislatures and Congress that are designed to promote recycling indirectly by affecting the markets for recyclables. The present study does not develop the implications for economic freedom of this flight from market-directed processes, but emphasizes instead the effects on economic efficiency

Cato Journal, Vol. 12, No. 2 (Fall 1992). Copyright © Cato Institute. All rights reserved.

The author is Associate Professor of Economics at Gonzaga University and Visiting Fellow at Resources for the Future.

of direct government involvement in recycling programs. However, it may be noted in passing that the array of proposals, many of which have been implemented, include landfill taxes, surcharges on disposal of recyclable materials, state tax incentives and subsidies to recycling programs, mandated minimum recycled material content in products, virgin (nonrecycled) materials taxes, public education programs, packaging legislation—including mandated package-to-contents ratios, and government procurement regulations stipulating recycled materials content (Kovacs 1988).

The perception that landfilling is unacceptable on the scale that has historically been utilized provides the major impetus for policies intended to stimulate greater recycling. The next section presents and critiques these objections to landfilling, and finds them to be based on a number of factual misreadings and misrepresentations. Subsequent sections take up in order the costs of operating recycling programs, basic economic principles as applied to solid waste management, the political dimensions of the problem, and an alternative approach to landfill siting which may facilitate more rational collective choice in solid waste management.

Landfilling

The Perception

Proponents of recycling depict the landfilling option as a dismal one indeed. The general features of the case against landfilling (replete with the inevitable rhetoric) may be summarized as follows.

Currently about 160 million tons of solid waste per year is discarded into the municipal waste stream, an amount that has swollen by 80 percent since 1960. Thirteen percent of this is incinerated and another 11 percent is recycled, leaving around 120 million tons disposed of in landfills (National Solid Wastes Management Association [NSWMA] 1989, p. 1). One year's solid waste, loaded into trash trucks, could form a convoy encircling the globe more than six times (Environmental Protection Agency [EPA], January 1989, p. 3) or reach halfway to the moon (EPA, February 1989, p. 8). While our use-it-once throw away society generates this swelling deluge of solid waste, landfill capacity becomes progressively limited. As a result of environmental and other problems, almost two-thirds of the approximately 4,500 municipal landfills that are either now open or that will be built in the next decade will close by the year 2000 whether or not they are full (NSWMA 1989, p. 2). The costs of landfilling are understandably rising precipitously. In some areas landfill

capacity is so limited, and the tipping fees so high, that solid waste must be transported out of state.

Compounding the problem of land scarcity is the serious environmental contamination resulting from solid waste landfills. Groundwater and surface run-offs percolating through landfills produce a toxic leachate containing heavy metals, acids, and other contaminants that eventually enter aquifers and streams. The resulting pollution poses health threats through direct water use and agricultural product contamination, and endangers wildlife and fisheries. Toxic gasses are also released, including methane and volatile organic carbons, which under certain conditions of confinement can even result in landfill explosions. To the unseen dangers of landfills may be added more apparent and inevitable smells, noise, wind-blown waste, and truck traffic which are the unhappy lot of those unfortunate enough to be situated near a landfill.

The foregoing summary provides a fairly comprehensive explanation for the widespread opprobrium accorded the landfilling of solid waste. The seemingly daunting task is next undertaken of showing why this view is seriously in error.

An Alternative View

Rather than figuratively elongating our trash as is the practice of those who would elevate our awareness and concern, let us place it in a single place. If it were piled to a depth of 100 yards, considerably less than the Fresh Kills landfill on Staten Island, the annual national space requirement would be a square area two-thirds of a mile on a side. One thousand years' trash would require less than a 30 mile square. The volume could be further reduced by half with the use (as is sometimes present practice) of a three-stroke compactor. The area required would then comprise three one-hundredths of one percent of the three million square mile area of the contiguous United States—a prospect surely disquieting only to the extreme claustrophobe. Moreover, the disposal area does not represent land rendered unavailable for other uses: land occupied by closed and inert landfills can be—and is in fact now being—put to a multitude of alternative uses, including parks, airports and golf courses. As archeologists are well aware, past civilizations have built and lived upon their rubble and rubbish—sometimes layer upon layer of it—for thousands of years (Rathje 1989, p. 65). It would appear reasonable to conclude that if there is a waste disposal problem, it does not relate to the gross space requirements of landfilling.

The amount of solid waste generated annually will probably continue to increase in the future as it has in the past. The space require-

ments above could be adjusted appropriately upward if we had good estimates of future waste growth. Although long-term forecasting is a dubious exercise at best, and will not be attempted here, the recent past may provide some guidance. Between 1960 and 1970 municipal solid waste grew at a 3.2 percent annual rate; however, from 1970 to 1986 the growth declined to 1.7 percent, and the amount entering landfills during that period grew at only 1.0 percent annually.¹ This rate is far less than the growth rate of goods and services consumption, and is indicative of a “throw away” society which in fact throws out a progressively smaller share of its output with the passage of time.

The landfill closure rates frequently cited by recycling proponents amount to little more than a numerical strawman. The great majority of landfills are small, designed for about ten years of operation, so that about half of these close in every five-year period. Large landfills are relatively small in number but absorb a very high percentage of the total solid waste tonnage. The real issue is whether new landfills can be sited and opened as closures occur. As indicated below, new landfills will undoubtedly be much larger on average than those presently in existence, so the relevant consideration is new landfill *capacity* rather than number.

The potential environmental damage wrought by landfilling is a matter of understandable and legitimate concern. However, a sharp distinction should be made between the concept of a modern sanitary landfill on the one hand and the traditional town dump on the other. The nation is peppered with the latter, some of which have leaked potentially toxic or hazardous substances into the ambient groundwater. However, the EPA's own analysis (EPA 1988) suggests that over 80 percent of landfills pose a very small lifetime health threat—less than one in one million under a worst case scenario. Even in locations where the potential for environmental contamination is a concern, it is feasible to construct and operate modern landfills which meet stringent environmental standards. Pursuant to Subtitle D of the Resource Conservation and Recovery Act of 1976, the Environmental Protection Agency in September 1991 enacted new landfill regulations to be effective October 1993 (see EPA 1991).² These regulations will be effective after October 1993, and apply to all landfills accepting municipal solid waste. A typical state-of-the-art landfill conform-

¹Calculated from data in Franklin Associates (1988).

²Rather than being performance criteria, the regulations tend to be in the form of technical standards, thus discouraging efforts to seek less costly alternative means of achieving the same results. However, some leeway is allowed in that EPA-approved state design and performance standards may be acceptable.

ing to the new standards will incorporate the following features (U.S. Government Printing Office, 1991):

- location restrictions related to wetlands, floodplains, fault areas, etc.,
- a liner system of clay and synthetic flexible membrane,
- controls limiting acceptance of toxic substances and sources,
- a leachate collection system,
- provision for transportation and treatment or disposal of leachate,
- a gas collection system,
- systems and procedures for materials handling, filling, gas and leachate monitoring, and surface water and groundwater control,
- daily earth covering,
- control of disease vector populations (rodents, flies, mosquitoes, etc.),
- restricted public access,
- site capping and closure procedures,
- monitoring and control of leachate, gas, and groundwater for several decades after closure.

An important part of the science of environmental engineering deals with the construction and operation of landfills of this type, many of which already exist in part as a result of state and local regulations (see Zandi 1989).³ Some have filled, closed, and as noted been converted to recreational or other uses, often to the accompaniment of escalating adjacent property values (NSWMA 1986).

Subtitle D standards will increase the scale economies of landfilling, making smaller sites less economical and resulting in fewer, larger landfills nationally. Higher quality landfills will entail higher costs, which will vary greatly, depending on a host of factors including land costs and climatic, soil and geologic conditions. The total cost per ton (including profit) for a landfill in the midwest which incorporates the features above has been estimated at less than \$20 per ton, as compared to about \$5 per ton (1988 dollars) for an acceptable landfill in 1975 (Glebs 1988).

It is enlightening to restate the cost from the perspective of the household. The landfill disposal cost of waste from a standard 32 gallon trash container, the contents of which weigh an average of about 21 pounds (Scumatz 1990, p. 12), would rise from about 5 cents

³Zandi (1989, p. 53) also observes that "there is no environmental objection to the landfill . . . that cannot for a cost be controlled. Landfilling of municipal solid waste [as distinct from industrial and hazardous wastes] is one human activity that can be designed such that all environmentally undesirable effects can be eliminated."

to 15–20 cents under this cost scenario. Households nationwide pay on average a considerably higher fee, equivalent to about 80 cents to \$1 per container, since most of the family garbage bill is for collection and hauling services. The total charge also frequently exceeds the total collection, hauling, and disposal costs, leaving municipalities with revenues from hidden taxes. In addition, explicit landfill taxes or special disposal fees are increasingly being instituted to subsidize recycling programs. It is also likely that returns in excess of costs (scarcity rents) are accruing to landfill owners due to the landfill siting difficulties discussed below. Finally, costs per ton are likely to be higher in areas where bi-weekly pickup is the norm. In view of the relatively large non-landfilling components of the total disposal cost, the effect of converting to the proposed EPA-standard landfills is to increase the average household's cost per container by only 10 to 15 percent.

In many cases it would be desirable, and perhaps possible (see footnote 2) to achieve compliance with reasonable environmental criteria without the necessity of many of the technical features listed above. Although landfilling costs will be commensurately higher in areas where land is costly, land prices have a surprisingly small effect on the total cost of waste disposal. For example, in deriving the \$20 per ton cost estimate just given above, Robert Glebs assumed a land price of \$8,000 per acre. This yields a total land cost which comprises only one-fiftieth of the total landfill cost. Recalling that landfilling cost is in turn a fraction of the total cost (including collection and hauling) of waste disposal, the portion due to land costs is even less. Even if landfilling comprised as much as one-third of the total cost, in the estimate above land would constitute only about *one one-hundred-fiftieth* of the total waste disposal cost including collection, hauling, taxes, etc.

To summarize thus far, the so-called solid waste crisis can not properly be ascribed to problems of national land scarcity, environmental degradation, nor even to excessive cost. The view amplified below is that the essence of the problem is the collapse of effective collective decisionmaking in the siting of new landfills. At this point, it will be useful to first consider the costs of recycling as a major landfilling alternative.

Recycling Program Costs

Explicit Costs

Recycling programs are highly varied. Organizationally, they cover gradations between private and public, and between nonprofit and

for profit, with differing mixes of volunteer and nonvolunteer labor and management. To differing degrees direct or indirect subsidies are provided. Although a typical program is usually based on either curbside pickup or drop-off centers, a combination of the two collection methods is usually employed. Some programs only consist of publicly funded collection and hauling, with separation, processing, storage and marketing left to outside private contractors; other programs encompass all of these activities. The range of materials recycled varies: aluminum cans, old newspapers and glass are standard items, but ferrous metal cans, yard wastes (for composting), corrugated containerboard, mixed papers, and other products may be included. Aside from variability of the programs, the types of the areas being serviced differ with respect to population density, residential type, quantity and composition of commercial and household waste, and public awareness and interest in materials recovery. Numerous factors, including the intensity of the public information and "education" effort lead to differences in administrative costs. Measures of program benefits also differ as a result of regional and temporal differences in recovered materials prices, as well as differences in avoided waste disposal costs.

The recycling literature is rife with descriptions and prescriptions relating to "successful" program operations. Information is plentiful on tonnages and percentages of solid waste recovered, numbers and percentages of households participating, and the like. Published program cost and revenue figures are rarely available, however, and even then are frequently unreliable or are in a form that is useless for purposes of evaluating true program costs and benefits. The patent intent of the presenters is usually the demonstration of a high level of program net benefits. Also, in principle certain adjustments should be made to the accounting costs and revenues in order to arrive at a closer estimate of net social costs and benefits. As examples, subsidies are not social benefits of programs, and should be excluded from revenues; avoided waste disposal "costs" should exclude the explicit or hidden disposal taxes alluded to earlier, as well as the avoided disposal cost of materials that would have been (or previously were) recycled even in the absence of the program.

Table 1 gives cost and revenue data on each of seven recycling programs reported on in California and New Jersey. The data are compiled from published reports from recycling officials in those states for periods of time in the early to mid-1980s. Aside from giving each program's size, no attempt is made here to go into descriptions of many characteristics of the different programs. The relevant comparison is between the net losses (inclusive of revenues on materials

TABLE 1
COSTS, REVENUES, AND AVOIDED SOLID WASTE DISPOSAL COSTS
PER TON FOR SEVEN RECYCLING PROGRAMS

Program Location/ Identification	Tons Recovered per Year	Program Cost per Ton	Revenues per Ton	Net Profit (Loss) per Ton	County Disposal Cost per Ton
<i>California Programs</i>					
Palo Alto	2,741	\$138.68	\$46.64	\$ (92.04)	\$47.00
Downey	876	101.72	18.79	(82.93)	34.00
Santa Rosa	2,570	62.90	21.97	(40.93)	44.00
Fresno	1,971	226.57	58.83	(167.74)	50.00
<i>New Jersey Programs^a</i>					
Program A	2,346	55.41 ^b	37.46	(17.95)	16.71
Program B	1,228	40.51 ^c	18.70	(21.81)	29.25
Program C	886	34.76 ^d	27.86	(6.89)	5.60

^a New Jersey recycling costs exclude land costs; disposal costs exclude collection and transportation.
^b Depreciation on capital estimated by 10-year straight line method.
^c Does not include depreciation on invested capital.
^d Does not include implicit value of volunteer labor.

SOURCE: Calculated from data in *The Solid Waste Handbook: A Practical Guide* (1986, pp. 239, 250-52).

sales) and the estimated avoided disposal costs. The California studies, which are far more complete, indicate that in three of the four cases reported recycling costs per ton are more than twice—in one case, more than three times—the avoided costs. For reasons obvious from the table notes, the New Jersey data can be taken as little more than roughly suggestive. There, one of the three programs shows net losses.

Information on other programs similarly indicates net losses on recycling programs. Even while referring to recycling as the “least cost option,” Rhode Island recycling officials report a program “net benefit,” including sales revenue and avoided collection and disposal costs, of *minus* \$40–50 per ton (Marks and Gold 1988). In reference to the City of New York recycling program, that city’s Commissioner of Sanitation has observed that “initially estimated by the Department to cost \$65 per ton, we now estimate that the collection and processing system we currently employ will cost between \$198 and \$273 a ton at full implementation” (Polan 1990, p. 4). High recycling program costs are endemic to communities large and small. Spokane, Washington, with a population of about 200 thousand, has embarked upon a recycling program which is costing over \$180 per ton, or four times the cost of collection, hauling and landfilling.⁴ It would appear reasonable to conclude on the basis of information provided by recycling officials themselves that in many cases the explicit costs of recycling programs exceed their benefits. As the following shows, when additional implicit costs are considered, the net costs of recycling are still greater.

Implicit Costs

The calculated costs of operating recycling programs exclude important components of their full economic or social cost. One is the implicit value of household time and effort, as well as the costs of other resources utilized by households in sorting, preparing, and possibly transporting recovered materials. Estimates of such costs should be added to the explicit costs of recycling to obtain the full cost. Some of these costs, although they may appear almost negligible when calculated per household on a daily basis, can add to surprising sums when totaled and aggregated. If only five minutes per week of household time enabled the recycling of 30 percent of solid waste now disposed of, then at \$8 per hour the household time cost would

⁴All recycling program costs given in this paragraph are net of materials sales revenues. The Spokane source is brochures and unpublished data provided to the author by Solid Waste Management, City of Spokane, 1991.

be about \$90 per ton. The total value of time used would total \$3 to \$3.5 billion annually. Adding to this the costs of household storage space and hot water for bottle washing would give a more complete picture of the unstated *household share* of the cost of recycling.

Although it may be objected that depriving people of small portions of leisure on an incidental basis is virtually costless, human behavior suggests otherwise. Commuters crowd and rush to avoid waiting a few minutes for a later subway train. Automobile drivers risk speeding tickets and express great dismay when traffic lights turn red at the last moment. The common use of devices such as radar detectors, electric can openers, remote TV and garage door controls, and automatic sprinkler systems also belie this hypothesis.

Upon reflection, the implicit cost figures above are not surprising. In being called upon to recycle, households are asked to make major changes in their lifestyles which will consume a portion of the valuable leisure that past material progress has made possible. Recycling, as a time-consuming activity, is most economically viable in times and places where wages and the foregone value of human time are low. In other places, metal cans are appropriately recycled to make roofs and used tires to make footwear. With sufficient maintenance, automobiles can be kept running almost indefinitely. However, efforts to force changes in this direction can be wasteful in a high income country such as the United States; and as per capita income grows the relative value of human time will increase and progressively raise the cost of recycling relative to landfilling and incineration. A given recycling percentage therefore becomes increasingly burdensome with the passage of time.

Noneconomic Rationals for Recycling

A fundamental objection to the manner in which recycling and landfilling costs have just been compared emanates from the view that certain resources possess inherent values that exceed their apparent market values. For example, to those by whom energy resources are regarded as inherently worth husbanding and withholding from use, waste disposal alternatives are more or less desirable according to their degree of energy utilization. In essence, this view asserts that it is desirable to use, say, \$100 worth of capital and labor to save \$50 worth of energy. Similarly, many proponents of recycling apparently favor the use of highly valued inputs of labor in order to avoid using lesser valued amounts of land or natural resources. To the extent that these views are based on other than economic considerations, an economic critique is understandably futile. However, it may be pointed out that such views are frequently

underpinned by a serious misreading of the facts. The perception discussed earlier that the nation is running out of land is but one of the misinformed justifications for the avoidance of natural resource use.

It is commonly believed that increased wastepaper recycling will save trees, and presumably result in a larger growing stock of forests in the long run. The adjustments that can be expected to occur in forest management cast serious doubt on this conclusion. About one-third of the pulpwood for paper comes from residues of other wood products, the production of which will be negligibly affected by recycling. The remainder mostly consists of pulpwood trees, largely plantation-grown softwoods planted in orderly rows and mechanically harvested as a 20-year rotation crop. Their small size—indicated by a fiber yield of less than 200 pounds per tree—contributes to the immense numbers of trees cited by recycling advocates as being “saved” from the woodsman’s axe.⁵ The notion that stately old trees are used for paper production is erroneous; their value as lumber or plywood far exceeds their pulpwood value. Increased recycling will result in the conversion to agricultural uses of some plantation forest lands in the same way that a reduction in the demand for bread will reduce wheatlands, the possible result being a net reduction in the nation’s forest inventory.⁶

We turn next to some general economic principles as they relate to appropriate levels of the different solid waste management alternatives.

Efficiency in Selecting Disposal Alternatives

There are four generally recognized ways of dealing with solid waste: (1) source reduction, i.e., changes in production and consumption patterns that lead to lower levels of waste generation, (2) recycling, (3) incineration, and (4) landfilling. The EPA has prioritized these methods into a hierarchy that places (1) and (2) above (3) and (4), but recognizes that “strict adherence to a rigid hierarchy is inappropriate for every community” (EPA February 1989, p. 17).

An economic approach to the selection of waste disposal alternatives begins with the explicit recognition that all methods of solid waste reduction are costly in the basic sense that each requires the diversion of resources that have alternative uses. Costs of landfilling and recycling are amply documented above. Source reduction is costly because it requires modification of production and consump-

⁵See, for example, County of Fairfax (1990, p. 1).

⁶See also Wiseman (1990, pp. 45–54).

tion behaviors which, given the choice, firms and households prefer. As examples, less packaging will lead to higher breakage, materials handling, spoilage and storage costs, and will increase spoilage-related waste disposal costs; household composting of yard waste uses time, land, and equipment.

Incineration and landfilling both entail associated direct costs due to resources used in the disposal processes. Other costs include the spillover or external costs of environmental degradation—for example, noise, air and water pollution—that are by definition not borne by either the waste facilities operator or his customers. The Subtitle D standards discussed above may be regarded as a means of internalizing these costs in the case of landfills. A full economic assessment of alternative forms of waste management would take such costs into consideration. It is further noted that external costs are not absent from the recycling alternative: a truck which loads and transports recyclables will generally contribute to noise, traffic congestion, and air pollution in much the same manner as one headed for a landfill. Just as with primary materials production, the processing of recycled materials inevitably leads to some release of toxic substances into the environment. De-inking of wastepaper, for example, results in releases of carbon tetrachloride, methylene chloride, tri- and tetrachloroethylene and numerous other chlorinated organics, and incineration of paper recycling residue emits dioxin and furan (Visali 1985, p. 242).

Given the costliness of all alternatives, a reasonable goal is the achievement of the best or least-cost (including external costs) method or combination of methods to employ in a given situation. In some cases a combination of source reduction, recycling, and landfilling may be optimal, in others landfilling alone may be optimal. The EPA hierarchy of alternatives presumably reflects the belief, despite the virtual absence of evidence, that the full marginal social and economic costs of landfilling exceed those of the other three options.

In principle, when two or more of the alternatives are used to deal with solid waste, costs will be minimized when each activity is carried up to a point such that the marginal or incremental costs are equal for each activity. For example, if recycling costs \$50 per additional ton and landfilling \$80, recycling activity should be increased and landfilling curtailed. With additional recycling, the costs per additional ton will rise above \$50, and eventually reach that of landfilling. The optimum mix of the two activities is reached where marginal costs are equated. Although practice may differ from this efficiency ideal, its recognition is fundamental to forming a ratio-

nal basis for the evaluation of waste disposal alternatives. The cost data above give evidence of recycling program *average* costs exceeding those of landfilling. It can be shown that this implies a still larger divergence of *marginal* costs, and hence that recycling programs are being operated at levels in excess of the optimum, with consequent waste of society's resources. The reasons for the various public policies which foster this waste are taken up next.

Sources of the Solid Waste Disposal Problem: Pricing and Siting

One of the factors underlying the problem of waste disposal is the failure of local governments to utilize weight- or volume-based pricing of solid waste collection and disposal services. In many parts of the country, notably the East, disposal charges to households consist of a flat-rate fee unrelated to the quantity of waste disposal services provided. As a result, the cost to a household of additional solid waste placed at curbside—the price of the service—is zero. Underpricing of solid waste disposal services encourages overuse of such services in two ways. By reducing the waste producer's incentive to consider disposal costs when acquiring materials that will require later disposal (including, for example, packaging and printed matter) the use and production of these materials is stimulated and underuse of the source reduction alternative results. For the same reason, underpricing of disposal also discourages voluntary recycling. Bringing the price of solid waste disposal into line with costs, including costs of the environmental safeguards discussed earlier, would allow households to make decisions based on the appropriate price signals, and would result in a lower demand for solid waste disposal services.

Because copious amounts of solid waste would be generated even under a more rational pricing system, the fundamental problem of landfill siting remains. The siting problem may be viewed as an example of a flaw in collective decisionmaking that has long been recognized in a number of other contexts. Policies having large per capita impacts on relatively small groups tend to elicit from those groups more active and effective political action than policies having small per capita effects which are diffused over larger groups. Reasons for this are that each member of the first type of group has enough at stake to induce individual action, and that the relatively small number and identifiability of the group members facilitates communication and joint action. Hence the larger group exerts proportionally less political influence, even in cases where a policy's

sum effect on the group exceeds that on the smaller group. For example, redundant military bases are kept in operation in response to local pressures, and tariffs and farm programs costly to consumers are enacted to protect specific groups of workers. The potential siting of a landfill usually creates a politically influential group of opponents which, anticipating the loss of amenities or property values, adopts the familiar not-in-my-backyard position. Despite the possible desirability of siting decisions for the community as a whole, local politicians find under the circumstances that either the granting of landfill permits or the decision to open a new municipal landfill is highly antithetical to astute reelection strategy. It should be emphasized that this problem has a very strong geographical dimension. In the more densely populated New England and Mid-Atlantic regions it is most acute, whereas in many parts of the West, with the exception of some major metropolitan areas, there is little difficulty in landfill siting.

Should sentiments lie strongly with the groups adversely affected by siting decisions, the following is offered for consideration: almost any economic change confers losses on some members of society. Expressway, school and prison construction has brought losses to affected property owners in the same way that handweavers, ironsmiths, and stable owners have suffered losses to the cause of economic change and development. Despite the losses, when such changes have been for the general betterment of society we have usually not shunned them, since doing so implies economic stagnation. Appropriate construction of landfills may be viewed in essentially the same light as these other examples, i.e., as a process under which societal gains exceed losses to the few.

Toward Landfill Siting

The political difficulties encountered in siting new landfills is the crux of the municipal solid waste disposal problem. In response, several proposals have been offered which would institute novel public decision processes to facilitate siting of landfills or other disposal facilities. For example, Robert Mitchell and Richard Carson (1986) describe a system incorporating mandatory local referenda in response to facilities siting applications. Howard Kunreuther and Paul Kleindorfer (1986) propose a sealed low-bid auction procedure to determine facility siting as well as tax redistributions among the affected communities. It may be parenthetically observed here that as long as environmental controls are adequate or environmental liability is clearly assigned and enforceable, landfill permitting

appears to be no more justifiable than the zoning of land for any other purpose.

Procedures along the lines of those just noted have in common two features. They both remove the permitting decision from local elected officials and include a mechanism for compensation to the host community. An interesting procedure which incorporates these characteristics has been utilized in Wisconsin for almost a decade. Faced with severe difficulties in landfill siting due to court support of local prohibitions, the state of Wisconsin in 1981 legislated procedures which facilitate permit issuance while still preserving environmental controls.⁷ The legislative intent was clear:

It is the intent of the legislature to create and maintain an effective and comprehensive policy of negotiation and arbitration between the applicant for a license to establish either a solid waste disposal facility or a hazardous waste treatment, storage or disposal facility and a committee representing the affected municipalities to assure that:

(a) *Arbitrary or discriminatory policies and actions of local governments which obstruct the establishment of solid waste disposal facilities and hazardous waste facilities can be set aside.*

(b) *The legitimate concerns of nearby residents and affected municipalities can be expressed in a public forum, negotiated and, if need be, arbitrated with the applicant in a fair manner and reduced to a written document that is legally binding.*

(c) *An adequate mechanism exists under state law to assure the establishment of environmentally sound and economically viable solid waste disposal facilities and hazardous waste facilities [Wisconsin Statutes, chap. 144; italics added].*

The essential feature of the Wisconsin law is its preemption of the power of local elected officials to decide whether or not to issue landfill permits. Municipal and/or county governments wishing to have an input to the permitting process are required to establish local negotiating committees in response to permit applications. The committees, which include a prescribed number of nonelected representatives, are empowered to negotiate the financial and other contractual relations between the landfill applicant on the one hand and affected residents and municipalities on the other. Examples of typically negotiated items include compensation to neighboring landowners, direct payments and disposal privileges of affected municipalities, traffic plans and road maintenance, hours of operation, fencing, arrangements for fire protection, and maintenance of environmental liability insurance. In general, environmental and

⁷For discussions, see Ruud and Werner (1985) and Schuff (n.d.).

technical matters are handled by a separate and parallel permitting process at the state level. Failure to reach a negotiated agreement at the local level can result in outside mediation, and should that fail, arbitration by a state agency. Perhaps it is the specter of the last that has resulted in only one arbitration of the approximately 120 permit applications that have been submitted since 1982.⁸

The workability of a system along these lines results from the explicit recognition and replacement of a flawed governmental decision process with a prescribed set of procedural rules. Although local elected officials are thereby constrained in their decisionmaking powers, they should have little cause for complaint, since their longevity in office can only be enhanced by the inability to make "unpopular" decisions.

Conclusion

Modern landfills are both an environmentally sound and a cost-effective means of disposing of municipal solid waste. A certain amount of recycling is also desirable, and a rational societal goal is the achievement of the right balance between these and other methods of solid waste management. However, the political difficulties attendant to the landfilling alternative have thrust public policies headlong toward a massive and costly national recycling effort. There is ample evidence that recycling is being pushed beyond the economically efficient level as a result not only of the siting problem but also because of misperceptions of the environmental impact of landfills, overestimation of the benefits of recycling, and underestimation of the real costs of recycling which importantly include household time and effort.

Efficient and rational management of municipal solid waste would be enhanced by three basic policies. One of these is the maintenance of adequate environmental standards for landfills. Albeit often mistaken, the perception that landfills are environmentally unacceptable has helped fuel the recycling engine, with the paradoxical effect of thwarting the development of cost effective and environmentally sound landfills. Second, pricing solid waste disposal services on a per unit basis at a rate equal to the full disposal cost would send a signal to users who could then adopt the amounts of recycling and source reduction that are consistent with true costs and benefits, including the value of household time. Finally, the basic source of the so-called solid waste crisis—the political difficulty of siting new

⁸Information provided to the author by officials of State of Wisconsin Waste Facility Siting Board.

landfills—is surmountable by reducing the power of local elected officials to grant or deny landfill permits. The legislation in force in the state of Wisconsin provides a model for circumventing the undesirable impact of special interest obstruction in the landfill permitting process. Although the Wisconsin procedure may be inferior in theory to auction approaches or even non-zoning of landfills, and cannot be expected to completely dissipate the misinformation and fervor which impels the frenetic national recycling movement, it does offer a politically practicable approach. If adopted elsewhere this approach could substantially remove the major impediment to the best use of resources in solid waste management and thereby relieve the mounting pressures for wasteful, governmentally supported recycling efforts.

References

- County of Fairfax, Virginia. Department of Public Works, Division of Solid Waste. "Newspaper Recycling Program: It's Picking Up." *Recycling News* (Spring 1990).
- Environmental Protection Agency, Office of Solid Waste. "EPA Issues Comprehensive Landfill Regulations." *Reusable News* (EPA/530-SW-91-022) (1991): 1.
- Environmental Protection Agency, Office of Solid Waste. *Recycling Works! State and Local Solutions to Solid Waste Management Problems*. Washington, D.C.: Government Printing Office, 1989.
- Environmental Protection Agency, Office of Solid Waste. *The Solid Waste Dilemma: An Agenda for Action*. Report. Washington, D.C.: Government Printing Office, 1989.
- Environmental Protection Agency, Office of Solid Waste. *Regulatory Impact Analysis: Proposed Revision to Subtitle D Criteria*. Report. Washington, D.C.: Government Printing Office, 1988.
- Franklin Associates, Ltd. *Characterization of Municipal Solid Waste in the United States, 1960 to 2000*. Final report prepared for Environmental Protection Agency, 30 March 1988.
- Glebs, Robert T. "Landfill Costs Continue to Rise." *Waste Age* (March 1988): 84-93.
- Glebs, Robert T. "Subtitle D: How Will it Affect Landfills?" *Waste Alternatives* (December 1988): 56-64.
- Kovacs, William L. "The Coming Era of Conservation and Industrial Utilization of Recyclable Materials." *Ecology Law Quarterly* 15 (4) (1988): 537-625.
- Kunreuther, Howard, and Kleindorfer, Paul R. "A Sealed-Bid Auction Mechanism for Siting Noxious Facilities." *American Economic Review* 76 (2) (1986): 295-99.
- Marks, Adam, and Gold, Marion. "Rhode Island Tackles Curbside Recycling." *Waste Alternatives* (June 1988): 34-38.

- Mitchell, Robert Cameron, and Carson, Richard T. "Property Rights, Protest, and the Siting of Hazardous Waste Facilities." *American Economic Review* 76 (2) (1986): 285-90.
- National Solid Wastes Management Association. *Landfill Capacity in the Year 2000*. Special Report, 1989.
- National Solid Wastes Management Association. *The Impact on Property Values of Solid Waste Facilities*. Technical Bulletin, 1986.
- Polan, Steven M. "Letter to the Honorable David N. Dinkins, et.al." 9 October 1990. Printed in Department of Sanitation, City of New York, *New York City Recycles, Preliminary Recycling Plan, Fiscal Year 1991, 1990*.
- Rathje, William L. "The Three Faces of Garbage—Measurements, Perceptions, Behaviors." *Journal of Resource Management and Technology* 17 (April 1989): 61-65.
- Ruud, Peter J., and Werner, Dean M. "Wisconsin's Landfill Negotiation Arbitration Statute." *Wisconsin Bar Bulletin* (November 1985): 17-19, 64, 65.
- Schuff, Richard G. "Solid Waste Siting in Wisconsin: An Effective Process." Mimeograph. Bureau of Solid Waste Management, Wisconsin Department of Natural Resources, n.d.
- Scumatz, Laura A. "Volume-Based Rates in Solid Waste: The Seattle Experience." Mimeograph. Seattle Solid Waste Utility, 1990.
- The Solid Waste Handbook: A Practical Guide*. New York: John Wiley and Sons, 1986.
- U.S. Government Printing Office. "Solid Waste Disposal Facility Criteria." *Federal Register* (9 October 1991): 50978-51119.
- Visalli, Joseph R. "Environmental Impact Considerations in Recycling Solid Wastes." *Journal of Resource Management and Technology* 14 (December 1985): 241-46.
- Wisconsin Statutes, Chap. 144, Sec. 144.445 (2).
- Wiseman, A. Clark. "U.S. Wastepaper Recycling Policies: Issues and Effects." Discussion Paper ENR 90-14, Resources for the Future, Washington, D.C., 1990.
- Zandi, Iraj, "Solid Waste Management Strategies—A Question of Priorities." *Journal of Resource Management and Technology* 17 (April 1989): 53-60.

WHY HAS THE LABOR-MANAGED FIRM FAILED

Svetozar Pejovich

Introduction

The end of the totalitarian socialist rule has created an institutional vacuum in Eastern Europe and the former Soviet Union. The transition process we are witnessing in that region is, in effect, the search for a new set of institutions.¹ We still do not know as much about the process of institutional change itself, about who should change the rules and under what circumstances the rules ought to be changed.² The transition process in Eastern Europe is, therefore, a gift from heaven for social engineers. And they have responded to this gift by flooding the intellectual and political markets with models, ideas, and proposals for the development of new institutional arrangements in the region.

Industrial democracy is an umbrella for all the different forms of labor participation in the management of business firms. A striking feature of industrial democracy is that it has failed to emerge spontaneously. It has also failed to perform successfully whenever and wherever imposed by fiat. Codetermination in Germany, the labor-managed economy in former Yugoslavia, and various labor participatory schemes in some West European and South American countries are good examples.

Cato Journal, Vol. 12, No. 2 (Fall 1992). Copyright © Cato Institute. All rights reserved.

The author is Professor of Economics at Texas A & M University. He wishes to thank Louis De Alessi, Enrico Colombatto, Joze Mencinger, John H. Moore, and Jim Dorn for their helpful comments. He also wishes to thank the Earhart Foundation and the Lynde and Harry Bradley Foundation for supporting his research on the economics of property rights in Eastern Europe.

¹Institutions are usually defined as the legal, regulatory, and custom-made arrangements for repeated human interactions—that is, institutions are the rules of the game whose major function is the predictability of behavior.

²Douglass North and members of the Public Choice School have contributed most to our understanding of the process of institutional change.

It would be, however, a mistake to underestimate the survival power of the concept of industrial democracy. Labor participation in the management of business enterprises offers too many opportunities for social engineering and satisfies too many ideological preferences to be discarded on account of its poor real world performance. This is especially true in Eastern Europe and the former Soviet Union where the end of the communist rule could not and indeed did not put an end to all the communist institutions and legacies. Thus, the search for new rules of the game in that region provides the supporters of industrial democracy with an opportunity to be heard (Vanek 1990). And their chances to be heard are likely to grow as free-market policies fail to produce quick returns.

This paper concentrates on a specific type of industrial democracy: *the labor-managed firm*. The support for the labor-managed firm comes from several sources: (1) the old communitarian tradition in most East European countries; (2) the collectivist mode of thinking in that region—a legacy of the communist rule; (3) labor leaders and social democrats in Western Europe; and (4) a handful of academic supporters in the West, such as Jacques Dreze, Henryk Flakierski, James Meade, Jan Svejnar, and Jaroslav Vanek. Also, some current proposals for privatizing state firms in Russia, Serbia, Slovenia, and Poland contain strong elements of labor-participation in management.

In a free society people get what they want to pay for. And that applies to all the different methods for organizing production. Investors in private-property, free-market economies are free to write any kind of agreement with their contractual partners. Indeed, we observe a large number of different types of business firms in the West. All those firms have emerged voluntarily and survived competition from other types of business organizations.

Dreze (1976), Meade (1974), Prasnikar and Svejnar (1990), and Vanek (1990), among others, have asserted that the labor-managed firm is or could be an efficient method of organizing production. Yet, this type of business organization has failed to appear on a significant scale in free-market, private-property economies. An implication is that the value of labor participation to the employees of business firms is worth less than the costs to their contractual partners of providing it. By failing to specify the property and contracting rights within which the labor-managed firm exists, Vanek and others have ignored the effects of negative incentives and high transaction costs on its ability to survive in competitive markets.

The purpose of this paper is to show how the bundle of property rights in the labor-managed firm creates two behavioral variables

that are *specific* to that bundle of rights (and of no significance in the private-property, free-market economy). Those two variables are the key to explaining the inefficiency of investment decisions by labor-managed firms.

Property Rights and the Labor-Managed Firm

The Yugoslav experience and academic research have shown that, like all other type of business enterprises, labor-managed firms could differ from each other in terms of their internal rules and organizations. However, the bundle of property rights within which any type of organization exists is the key to analyzing how it functions. It is, therefore, necessary to identify the basic bundle of rights which sets the labor-managed firm apart from other types of business enterprises. This bundle of property rights is characterized by the following five features:

1. *The employees govern the firm.* The employees are decision-makers in their respective firms. We assume that the firm's decisionmaking structure, chosen by either its employees or the state (on their behalf), takes into account preferences of the median worker.
2. *The employees have claims on the firm's cash flows.* This right says that the employees of the labor-managed firm are responsible for all the financial and legal obligations of the firm as well as the allocation of the firm's residual to any purpose (i.e., the wage fund, investment fund, reserve fund) that is not explicitly forbidden by law.
3. *The employees' rights specified under (1) and (2) are not transferable.* The right of ownership³ in (1) and (2) would, in effect, allow the employees to sell both their right to govern the firm as well as their residual claims on the expected future cash flows of the firm. The labor-managed firm would then turn into a private-ownership firm.
4. *The employees' rights specified under (1) and (2) are contingent on their employment with that firm.* A worker's claim on his share of the residual terminates when he leaves the firm. To extend his non-tradeable claims on the firm's cash flows beyond that time would be costly to implement and monitor. It would also impose a rising burden on the firm's future costs of production. This specific right makes the employees' time horizon (the

³Transferability of assets at mutually agreed prices is a critical component of the right of ownership.

expected length of stay with their firm) a critical variable in the labor-managed firm's choice of investment projects.

According to Eirik Furubotn (1976), the time horizon of the employees of the labor-managed firm is rather short. It is also likely to fall short of the employees expected termination dates. First, an important implication of the rights specified under (2) and (3) above is that the employees cannot diversify their portfolio of non-tradeable claims on the firm's cash flows. Thus, we should expect to observe a conservative bias in the firm's decisions that have future consequences. An important paper on the codetermining firm in Germany (Benelli, Loderer, and Lys 1987) confirms this proposition. Second, hiring younger workers could easily shrink the employees' time horizon. A 45-year-old worker has, in general, a shorter work horizon than a younger man. However, his time horizon with the firm is likely to be longer than that of a younger man who hopes to move up by moving around.

5. *The labor-managed firm has no ownership in its capital assets.* The employees determine the use of capital assets held by their firm; however, they only hold non-tradeable claims on the firm's cash flows while they work for that firm.

Research has shown that the bundle of rights specified under (1) - (5) above creates some negative incentives and positive transaction costs that are responsible for the inefficiency of investment decisions by labor-managed firms (Jensen and Meckling 1979, Milovanovich 1990, Pejovich 1990). To alleviate those effects, some writers have proposed various immunizing stratagems (Flakierski 1989, pp. 67-70) which either ignore transaction costs of making, maintaining and enforcing those changes, or tend to privatize labor-managed firms, or both.

The Labor-Managed Firm and Investment Decisions

The labor-managed firm can acquire capital assets through a state agency, by renting them from others, and by issuing debt claims. Also, the employees can allocate a part of the residual into the firm's investment fund.

Acquiring capital through a state agency has many efficiency implications. It takes an act of faith to assume that the social welfare function exists, that the state knows it, that the state is going to do something about it, and that the state will be willing to relinquish its control over the assets given to enterprises. There is no empirical

evidence to support this type of naive expectations about the behavior of any government (Brunner 1987).

The rental option for acquiring capital assets has two major problems. First, a number of intangible productive assets, such as the firm's investment in the distribution systems, the design of products, and training of the labor force, cannot be rented. Second, the rental of durable productive assets is a costly method of acquiring capital assets.⁴

Given its bundle of property rights, the labor-managed firm's main sources for acquiring capital assets are then the claims of debt-owners, and the personal contributions of the employees. The former are pure financial claims. As for the latter, since the employees have no ownership claims on the firm's capital assets, they cannot have claims on the monies they give to the firm to purchase new assets. The employees only hold non-tradeable claims on the firm's return from those assets, and even that for only as long as they stay with the firm.

External Financing of Investments by the Labor-Managed Firm

The sources of external funds could be state and/or bank loans, sale of bonds, interfirm borrowing, and perhaps some other monies. This paper concentrates on the effects of bank credit on investment decisions by labor-managed firms. Other external sources of funds would be subject to similar incentives.

The employees' benefits from investments financed by bank loans. The employees of the labor-managed firm have non-transferable claims on the stream of annual returns (B) from any specific investment (I) made by their firms. This property right creates incentives for the employees to transfer the firm's cash flows from the future to the present. The employees can run down inventories, fail to replace capital assets, under-invest in the maintenance of capital goods, vote themselves large pension benefits with no funding provisions, grant themselves large severance payments, sell long-term bonds with no sinking-fund provisions, and so on. To alleviate the effects of those incentives, the government has to invest resources in enacting, maintaining, and enforcing a number of constraints on the property and contracting rights of the employees. Examples are

⁴The obvious agency costs of the rental arrangement are those associated with the reduced incentives for the user to maintain the assets properly, to guard it from theft, and the increased incentives to misuse it. ...The magnitude of these costs along with the monitoring and bonding costs that would be incurred in the effort to control them explains why rental or leasing of most durable production goods is not observed. It is simply a more costly contracting arrangement (Jensen and Meckling 1979, p. 480).

the depreciation rules for capital assets, the rules for maintaining and repairing physical assets, the rules for severance payments, and so on. An important consequence of those constraints is that they require a costly bureaucracy that is *specific* to the labor-managed firm.⁵

Two implications of the employees' non-tradeable claims on the firm's cash flows are the following. (1) The employees do not view equivalent-present value projects as being equal—those projects whose returns occur more quickly are preferred to those with even flows, and the latter are preferred to investments whose yields are bunched in later periods. (2) The absence of financial markets means that the rate of interest does not express the present prices of capital goods relative to their current costs of production; it merely measures a cost of investment.

An employee's benefits from any specific investment (I) are limited to the annual returns (B) from that investment over the employee's expected stay with the firm (t).⁶ The time horizon of the median member of the collective or whatever the decisionmaking group happens to be is then a critical factor in choosing investment projects in the labor-managed firm. In comparison, the time horizon of an investor in the private-property, free-market economy is irrelevant because the flow of benefits over the productive life of his assets is available to him in one lump sum.

The employees' costs from investments financed by bank loans. The employees' cost of any specific investment (I) financed by bank credit is the series of payments (C) to the bank over the employee's time horizon (t). Given the rate of interest, the firm's annual payments

⁵That is exactly what happened in Yugoslavia, the only country that has experimented with this system on a large scale and over a period of time long enough to provide us with a strong data base. Prasn timer and Svejnar (1990, p. 5) missed the point when they tried to separate a theoretical construct of the labor-managed firm from the Yugoslav experience: The strong influence of the League of Communists, the Communist-oriented trade unions and the various government authorities suggests that one ought to examine seriously the extent to which the behavior of Yugoslav firms resembles that of a proto-typical socialist enterprise rather than that of a labor-managed firm. The labor-managed firm needs the political monopoly to protect it from competing methods of organizing production, and the economy needs the self-management bureaucracy to protect it from the employees' incentives to eat up their firms.

⁶The employees' benefits could also be expressed in terms of their present values as follows:

$$PV = [B(1+i)^t - I]/i(1+i)^t$$

where PV is the present value of the flow of returns (B) from a specific investment (I) over the employees' time horizon (t). PV equals the true present value only when (t) is equal to or greater than the expected life of that investment. This paper focuses on annual returns and costs. It is a simpler approach that has no effect on our results.

depend on the length of time over which the loan has to be repaid (n). The length of bank loan is an item in the contract that is negotiated between the borrower and the lender. It is then important to identify the borrower's and the lender's incentives with respect to (n).

Suppose the labor-managed firm secures an investment loan that has to be paid back to the lending bank over a period of time (n). If the time horizon of the firm's employees (t) were equal to or greater than the length of that loan (n), the current group of workers would bear the entire costs of investment (I). However, *if the employees' time horizon fell short of the length of loan, a part of the total cost of investment (I) would be shifted to the next generation of workers in the same firm.*

The employees of the labor-managed firm have then incentives to seek investment loans with as lengthy repayment schedules as the banks are willing to go along with, and to negotiate agreements to pay only interest on investment loans over the current employees' time horizon.

The employees' annual costs of any specific investment (I) financed by bank credit is then

$$C = I \frac{i(1+i)^n}{(1+i)^n - 1}, \quad (1)$$

where (n) is the length of bank loan.

The choice of investment projects by labor-managed firms. The employees' benefits from an investment project are the expected returns (B) over their time horizon. The employees' costs from that same investment are the annual payments to the bank (C) over their time horizon. It follows that the investment decision of the labor-managed firm depends on the relationship between these two flows. To simplify discussion, we assume that both flows are uniform over the employees' time horizon.⁷ Then,

$$B \geq I \frac{i(1+i)^n}{(1+i)^n - 1}. \quad (2)$$

Since the length of loan (n) is determined *contractually* between the lender and borrower of investment funds, it is possible to adjust (n) so that

$$B = I \frac{i(1+i)^{n^*}}{(1+i)^{n^*} - 1}. \quad (3)$$

⁷Uneven flows of (B) and (C) could generate different technical solutions but they would not change the effects of the bundle of property rights in the labor-managed firm on the employees' incentives in choosing investment projects.

The equality (3) shows the *minimum* (n^*) required to make a specific investment project acceptable to the employees of the labor-managed firm. The investment decision of the labor-managed firm then depends on the ability of its manager to negotiate a loan that, at the minimum, has to be repaid over a period of time that makes the flow of returns from an investment equal to its production costs times the flow of income over (n^*) years from \$1 now. *Whatever the rate of interest, the length of bank credit (n) is a significant, perhaps the most significant, variable in the firm's choice of investments.*⁸

In comparison, the decisionmaker's time horizon and the length of bank loan play no role in the choice of investment projects in the private-property, free-market economy.

Changes in the length of bank credit and their implications. Let us now look at the effects of changes in the length of bank loans on the efficiency of investment decisions by self-managed firms.

The employees of a labor-managed firm are looking at an asset that costs \$1,000,000, has a productive life of 15 years, and promises to yield \$118,326 per year over a period of 15 years. At the market rate of interest of 10 percent, the present value of that asset is \$900,000, and the investment is clearly inefficient. However, suppose that the manager of the firm is able to borrow \$1,000,000 from the bank on a 20 year loan at 10 percent. A private-ownership firm would still turn the project down, but the labor-managed firm would not. The employees' annual payments to the bank would be \$117,460 per year, while their annual benefits are \$118,326. As long as the employees' time horizon is 15 years or less, the employees would have \$866 each and every year to divide among themselves. So they would make an inefficient investment. The costs shifted to the next generation of workers (including those current employees whose time horizon exceeds 15 years) would equal annual payments of \$117,460 per year over the last five years of the contract. If the employees' time horizon were, say, 12 years, the costs shifted to the next generation of workers would be \$117,460 per year for the remaining eight years, and so on.

If the length of loan is 19.57 years, the firm's annual payments (C) to the bank would be equal to annual returns (B) from that investment. However, a 20 year loan gives the employees an additional \$866 per year *over their time horizon* with the firm. It means that

⁸Our results are contrary to those Meade (1972), Vanek (1970, 1977), and other advocates of the labor-managed firm who consider the market (clearing) rate of interest to be a critical variable that assures the system of the efficiency of externally financed investments.

TABLE 1
RELATION BETWEEN THE LENGTH OF LOAN AND REWARDS
OFFERED

Length of Loan	Benefits (<i>B</i>)	Costs (<i>C</i>)	Benefits—Costs
19.56 years	\$118,326	\$118,326	\$ 0
20	118,326	117,460	866
21	118,326	115,624	2,702
22	118,326	114,005	4,321
23	118,326	112,572	5,754
24	118,326	111,300	7,026
25	118,326	110,168	8,158

the manager of the firm has incentives to offer some people at the bank an equivalent of up to \$866 per year in cash or specific goods in exchange for a 20 year loan. And he would have incentives to offer even better rewards for longer loans. Table 1 shows how changes in the length of loan increase the amount of money the manager can use to seek a good contract.

Suppose the time horizon of the employees is 10 years, and the firm has three investment projects under consideration. The productive life of each project and their respective annual streams of benefits per \$1 are shown in Table 2, columns (1) and (2). Column (3) shows the minimum length of bank loan (n^*) that makes the flow of annual

TABLE 2
PROPERTY RIGHTS AND THE INVESTMENT DECISION

Life of Asset	Benefits (<i>B</i>)	Minimum Length of Loan(n^*)
10 years	\$.146	12.12 years
15	.118	19.73
20	.106	30.13

benefits equal to the employees' annual costs from those investments. At a 10 percent rate of interest, a private-property firm would clearly turn all three projects down because the present value of each project at the 10 percent rate of interest is 90 cents. However, if the manager of the labor-managed firm were able to negotiate bank loans in excess of (n^*) years the employees would find all three investments acceptable.

Internal Financing of Investment by the Labor-Managed Firm

The employees' non-tradeable claims on the firm's cash flows create two wealth-increasing options for them. The employees can take the firm's residual out as wages and invest in assets they would own. The employees can also leave a part of the residual with the firm for investment in new capital assets in which they would hold non-tradeable claims contingent on employment by the firm.

Property rights differences between those two options have been researched in some detail (Pejovich 1969, Furubotn and Pejovich 1970, Jensen and Meckling 1979). The major results of that research are summarized here in order to complete our discussion on investment decisions of the labor-managed firm.

The employees have the right of ownership in the assets they purchase as individuals. They can trade those assets, give them away, will them to their heirs, or do anything else that is not explicitly forbidden by law. On the other hand, the employees have only non-tradeable claims on the firm's year-to-year cash flows contingent on employment. That is, they have claims on neither the amount of the residual they leave with the firm for purchase of new productive assets nor the year-to-year returns from those assets once they leave the firm.

The employees' cost of any specific investment (I) financed from the firm's residual is the sum total of annuities (y) they could have received over the period (t) by taking the amount equal to (I) out as income and saving their respective shares at (i) rate of interest. The annuity per \$1 is

$$y = \frac{i(1+i)^t}{(1+i)^t - 1} \quad (4)$$

The employee's benefits from the same investment (I) financed from the firm's residual is the flow of returns (B) over their time horizon (t).

For the employees to be indifferent between internally financed investments and their own individual savings, the former must earn the rate of return (r^*) equal to (y) in (4); that is, the required rate of return (r^*) for internally financed investments must be equal to the annuity from one dollar at the rate (i) over the employees' time horizon (t).⁹ The rate of return (r^*) is the rate of interest (i) *adjusted*

⁹The assumption is that the two investment alternatives are alike with respect to risk level and liquidity. It is a simplifying assumption that has no effect on the results of our analysis in the paper.

for the incentives effects of property rights in the labor-managed firm:

$$r^* = \frac{i(1+i)^t}{(1+i)^t - 1} \quad (5)$$

The rates of return that would make the employees indifferent between investments that are internally financed and their own private investments could be easily calculated. At a 10 percent interest rate and time horizons of 1, 5, 10, and 15 years, they are 110 percent, 26 percent, 16 percent, and 13 percent, respectively.¹⁰ Given the firm's schedule of investment opportunities, the difference between the market rate of interest (i) and the required rate of return (r^*) must result in an inefficient level of self-financed investment by labor-managed firms.

Conclusion

This paper describes the prevailing property rights in the labor-managed firm, identifies incentives and transaction costs that are *specific* to that bundle of property rights, and analyzes the effects of those incentives and transaction costs on the employees' investment decisions. Reminding those who advocate labor participation in the management of business firms that institutions matter, Jensen and Meckling (1979, p. 480) wrote,

Ignoring the agency costs of alternative contractual forms in comparing two systems where the only difference between the two is the contractual form allowed is unlikely to shed light on the major issues. But Dreze, as well as most others writing on the topic, does exactly this.

The analysis of this paper shows that *the length of bank credit* (or any other financial claim) and *the employees' time horizon* are the two critical variables for the employees' choice of investments. And those two variables are specific to the bundle of property rights in the labor-managed firm.¹¹

With respect to externally financed investments, the analysis shows that when the length of bank credit exceeds the employees' time horizon and the employees' time horizon falls short of the pro-

¹⁰As in the case of externally financed investments, equivalent-return projects which pay off quickly would be favored relative to those whose payoffs occur later in the future.

¹¹When the employees' time horizon is greater than the productive life of assets, the relationship between (t) and (n) would determine whether an inefficient investment is chosen.

ductive life of assets, the labor-managed firm has incentives to finance inefficient investment projects. The fact that the labor-managed firm has not emerged voluntarily on a significant scale in private-property, free-market economies is evidence that the market for business organizations considers the length of bank credit to be greater than the employees' time horizon, the employees' time horizon to be shorter than the productive life of assets, and the lender's incentives insufficient to offset the employees' incentives to seek loans in excess of their time horizon.

The labor-managed firm would avoid making inefficient investments *only if* the prevailing incentives and transaction costs pushed the employees toward equalizing the length of bank credits with the expected life of capital goods to be purchased with those credits. For that to happen, several conditions have to be satisfied. First, the employees of the labor-managed firm must have incentives to negotiate contracts that would equalize the length of bank credits with the life of assets to be purchased with those credits. It is clear from the analysis in this paper that the employees *do not have* such incentives. Second, the cost of information to the lender about the expected life of capital goods, including intangible assets, must be low relative to that of the borrower. However, the absence of tradeable claims in financial markets raises the lender's information costs about both the productivity as well as the life expectancy of capital goods to be financed by loans. Finally, the lender must have strong incentives to incur the cost of negotiating contracts that would equalize the length of his loans with the life of assets to be purchased with those loans. There are many possible types of lending institutions, such as private-ownership banks, labor-managed banks, various financial institutions, government agencies, and specialized super banks. The analysis of all those types of lenders and their incentives is beyond the scope of this paper. It is, however, difficult to envisage a bundle of rights in all those lending institutions that would create incentives and transaction costs sufficient to change this paper's conclusions.

References

- Benelli, Giuseppe; Loderer, Claudio; and Lys, Thomas. "Labor Participation in Corporate Policy-Making Decisions: West Germany's Experience with Codetermination." *Journal of Business* 60 (1987): 553-75.
- Brunner, Karl. The Limits of Economic Policy. In *Socialism: Institutional, Philosophical and Economic Issues*, pp. 33-54. Edited by Svetozar Pejovich. Dordrecht: Kluwer Academic Publishers, 1987.

- Dreze, Jacques. "Some Theories of Labor Management and Participation." *Econometrica* 44 (November 1976): 1125–39.
- Flakierski, Henryk. *The Economic System and Income Redistribution in Yugoslavia*. Armonk, N.Y.: M.E. Sharpe, 1989.
- Furubotn, Eirik. "The Long-Run Analysis of the Labor-Managed Firm: An Alternative Interpretation." *American Economic Review* 66 (March 1976): 104–23.
- Furubotn, Eirik, and Pejovich, Svetozar. "Property Rights and the Behavior of the Firm in a Socialist State." *Zeitschrift für Nationalökonomie* 30 (1970): 431–54.
- Jensen, Michael, and Meckling, William. "Rights and Production Functions: An Application to Labor-Managed Firms and Codetermination." *Journal of Business* 52 (1979): 469–506.
- Meade, James. "The Labor-Managed Firm in Conditions of Imperfect Competition." *Economic Journal* 84 (December 1974): 817–24.
- Milovanovich, Milic. *Teorija i Praksa Samoupravljanja*. Beograd: Naucna Knjiga, 1990.
- Pejovich, Svetozar. "The Firm, Monetary Policy and Property Rights." *Western Economic Journal* 7 (September 1969): 193–99.
- Pejovich, Svetozar. "A Property Rights Analysis of the Yugoslav Miracle." *Annals* 507 (January 1990): 123–32.
- Pejovich, Svetozar. "A Property Rights Analysis of the Inefficiency of Investment Decisions by Labor-Managed Firms." *Journal of Institutional and Theoretical Economics* 148 (March 1992): 30–41.
- Prasnikar, Joze, and Svejnar, John. "Workers' Participation in Management vs. Social Ownership and Government Policies." Working Paper no. 264, Department of Economics, University of Pittsburgh, 1990.
- Vanek, Jaroslav. *The General Theory of the Labor-Managed Economy*. Ithaca: Cornell University Press, 1970.
- Vanek, Jaroslav. *The Labor-Managed Economy*. Ithaca: Cornell University Press, 1977.
- Vanek, Jaroslav. "Beware of the Yeast of the Pharisees." *Ekonomika Analiza* 24 (1990): 113–24.

COMMUNICATIONS

THE SEARCH FOR MACROECONOMIC STABILITY: COMMENT ON SUMNER

W. William Woolsey

In a recent issue of the *Cato Journal*, Scott Sumner (1991) discussed the development of proposals to stabilize economic aggregates. He described early proposals to stabilize the price level by using a compensated dollar and a problem with those schemes due to an "information lag." He proposed a solution to that problem using index future convertibility and discussed alternative targets that have been developed by modern macroeconomists.

Sumner briefly discussed the BFH system of free banking and claimed that it too would be subject to difficulties caused by the information lag. While his specific argument fails because he misunderstood BFH, the information lag could create a serious problem. Fortunately, that problem could be solved by combining BFH with his proposal for index future convertibility.

Sumner on the BFH System of Free Banking

Sumner's discussion of the BFH system is marred by two small errors and one serious misunderstanding. First, Robert L. Greenfield and Leland B. Yeager named the BFH system to credit ideas borrowed from Fischer Black (1970), Eugene Fama (1980), and Robert Hall (1982). But BFH was not their name for the various proposals of Black, Fama, and Hall (Sumner 1991, p. 752). It is a different, though related, proposal (Greenfield and Yeager 1983, p. 302). Further, Sumner emphasized proposals and cited papers (such as Black 1981, Fama 1983) different from those emphasized by Greenfield and Yeager.

Second, Greenfield and Yeager never described the BFH dollar as being abstract (Sumner 1991, p. 752). An abstract unit of account

Cato Journal, Vol. 12, No. 2 (Fall 1992). Copyright © Cato Institute. All rights reserved.

The author is Associate Professor of Economics at The Citadel. He wishes to thank Scott Sumner and an anonymous referee for helpful comments and suggestions.

would be *undefined*; independent of any medium of account. The BFH dollar would be defined, but not as a unit of money. In contrast to a conventional monetary system, a nearly comprehensive bundle of goods and services would serve as medium of account (Greenfield and Yeager 1983, p. 303). (While dollar-denominated checks, deposits, banknotes, and token coins would surely exist in BFH, none of those items would serve to define the dollar.)

How BFH Differs from the Compensated Dollar

More serious than those small errors was Sumner's misunderstanding of indirect convertibility. He apparently confused BFH with a privatized version of Fisher's compensated dollar. In the BFH system, dollar-denominated checks, deposits, banknotes, and token coins would be redeemed with an amount of gold (or other agreed redemption medium) having a market value equal to that of the bundle of goods defining the dollar (Yeager 1985; Greenfield and Yeager 1986, 1989). That institution is significantly different from a compensated dollar: redeemability with gold at a temporarily fixed price that is adjusted periodically in response to changes in a price index.

In a footnote (1991, p. 752), Sumner criticized Greenfield and Yeager (1989, p. 419) for stating that the dollar's gold content should be adjusted in view of changes in the price of gold. He argued that adjustments in the gold content of the dollar based on incipient changes in the price level should be sufficient to provide price level stability. His argument would be correct if BFH were a compensated dollar that utilized extremely rapid (or even preemptive) adjustments to the price of gold. According to his understanding of BFH, a change in golds' supply or demand conditions would imply redemptions (or deposits) at its temporarily fixed price, a change in the quantity of money, and a change in the price level. That incipient change in the price level would prompt a preemptive change in the price of gold. Price level stability would, therefore, be maintained.

Greenfield and Yeager's statement, however, suggests that a quite different market process applies to BFH. (The process outlined below is described in more detail in Woolsey and Yeager 1992.) Unlike Fisher's compensated dollar (or a conventional gold standard), the BFH system would not interfere with the normal workings of the gold market. There would be no official price of gold; it would not be fixed even temporarily. Assuming the sum of the market prices of the items in the bundle remained at the defined price of \$1, the banks would charge the going market price of gold to those redeeming

checks, deposits, notes, or coins. Both the price on the market and at the banks' redemption windows would respond directly to shifts in the supply or demand conditions for gold. BFH would *not* involve changing the gold content of the dollar (i.e., a targeted dollar price of gold) to pre-empt incipient changes in the price level.

The BFH system would allow incipient changes in the price level to be preempted, but without any change in the gold content of the dollar. First, suppose no such preemption occurred and some aggregate supply or demand shock caused the price of the bundle to *actually* rise above \$1. The banks' redemption obligation would compel them to sell gold at their redemption windows at a price that would be less than the current market price of gold in inverse proportion to the rise in the price of the bundle. Gold's market price would continue to adjust freely according to its supply and demand conditions, but a bank would be forced to provide gold to those redeeming its checks, deposits, notes, or coins at a slightly lower price. To stop the financial losses caused by buying gold on the market high and then selling it low, banks would sell bonds, contract loans, and raise their deposit interest rates. The consequent contraction in spending would return the price of the bundle to its defined price while simultaneously allowing the banks to again charge the market price of gold at their redemption windows.

Since the *difference* between gold's market price and its price at the banks' redemption windows would prompt the corrective deflationary forces, the gold content of the dollar would be unimportant. It could be changing simultaneously because of conditions peculiar to the gold market. Given the relative price of gold, however, selling gold for something other than its market price during any period when the price of the bundle was actually deviating from \$1 would amount to a temporary change in the gold content of the dollar at the banks' redemption windows.

Even temporary changes in the gold content of the dollar could be avoided, however. Banks would have an incentive to avoid financial losses entirely by adjusting credit conditions and deposit interest rates enough to reverse undesirable changes in spending before any changes in the price of the bundle actually occurred. Should banks correctly anticipate changes in the price of the bundle, inflationary or deflationary shocks would be offset without any change in dollar price of gold at the banks' redemption windows.

The BFH System and the Information Lag

Despite his misinterpretation of the BFH system, Sumner makes an important point. The BFH dollar would be defined relative to

something like the market basket used to measure the CPI, WPI, or GNP deflator. (Contrary to Sumner 1991, p. 752, that is *clearly* what Greenfield and Yeager had in mind. Otherwise, BFH would not stabilize the price level.) Sumner explains that an information lag exists because time is needed to measure macroeconomic variables (1991, p. 752). It seems likely, therefore, that the price of the bundle of goods and services defining the BFH dollar would be reported only periodically.

Sumner argues that the information lag would imply problems for BFH, but his rationale is wrong. He describes destabilizing speculation in the context of the compensated dollar (1991, p. 750) and implies the BFH system would suffer from the same difficulty. (He explicitly makes this claim in 1990, p. 116.) If the CPI was to rise above its target, that fact would be apparent to speculators just prior to the announcement. Since the compensated dollar would require the government to decrease the official price of gold in inverse proportion to the increase in the CPI, speculators would sell large amounts of gold to the government at the official, temporarily fixed price just prior to the announcement and then buy it back at a slightly lower price immediately after the announcement. The sale of gold to the government would cause an unwanted increase in the quantity of money. Further, the government's losses could be ruinous.

That type of speculation could not exist in the BFH system, since the market price of gold would adjust freely. If speculators were to sell gold for any reason, the price of gold on the market and at the banks' redemption windows would decrease immediately. There would be no consequence for the quantity of money and no speculative losses for the banks.

The information lag could cause a disaster in the BFH system for a different reason. If the price of the bundle actually rose above \$1, the banks would be obligated to sell gold for less than its market price. But for how long? It would be absurd for banks to suffer the consequent losses on all redemptions occurring between the announcement of a high price for the bundle and the subsequent announcement. In the context of monthly announcements, such a requirement would surely result in the collapse of the banking system.¹

¹The literature on BFH has assumed that continuous measurement of the price of the bundle is possible. In correspondence, however, Yeager has suggested that periodic measurements could be accommodated by redeeming checks, notes, and coins with an estimated amount of gold and then extrapolating between the previous and subsequent measure of the price of the bundle to determine the needed adjustment to the preliminary settlement. His suggestion does amount to a complicated form of index futures convertibility.

Index Future Convertibility

Sumner is proposing an intriguing solution to the problems created by the information lag. An index futures contract would be used as an instrument of monetary policy. He suggests that the Federal Reserve be required to buy and sell unlimited quantities of an index futures contract on the CPI at a target value of 100. (In his 1989 paper, he described a similar scheme using a target for nominal income.) If speculators expected the index to be above target, they would buy futures from the Fed. If they expected the index to be below target, they would sell futures to the Fed (Sumner 1991, p. 753).

Sumner's proposal requires some modification, since the purchase or sale of futures contracts would not be equivalent to ordinary open market operations. Index futures contracts are promises to make payments, so the Fed's trades would have no immediate consequence for the quantity of money. (If the Fed imposed cash margin requirements on the transactions, it would cause a contraction in the quantity of money both when it sold and when it bought futures contracts.)

Sumner (1991, p. 754) solves that problem by proposing a modification aimed at reducing the risk premium on the futures contract. He proposes that the Fed set low margin requirements and use parallel open market operations to avoid taking any significant long or short position. The ordinary open market operations would cause the changes in the quantity of base money needed to keep the expected value of the CPI on target.²

Index Future Convertibility and the BFH System

Index future convertibility could be applied to BFH, thereby avoiding any difficulties due to the information lag. To describe how such a scheme would operate, some detailed assumptions about the payments system must be made. Like all versions of BFH, there would be no government currency or other base money. Suppose banks offered checkable deposits and issued banknotes (notes) and token coins (coins). Those items would be denominated in dollars, but differentiated by issuer. The banks would accept each others'

²Sumner credits Hall (1983) for inspiring his scheme (1991, p. 752). David Glasner proposed a system (for which he credits Earl Thompson 1986) that is equivalent to index future convertibility (1989, pp. 230–36).

items for deposit at par, so checks, notes, and coins drawn on all banks would be generally accepted in exchange.³

The banks would participate in a clearinghouse. Suppose it provided each bank with a clearing account that began with a zero balance. As checks, notes, and coins cleared, the banks would develop net credit or debit balances. The clearinghouse would pay interest on net credit balances and charge interest to net debit balances. Suppose the clearinghouse made the rates it paid and charged unfavorable compared to those on overnight loans, giving banks an incentive to use overnight loans (or some other money market instrument) to offset any excessive or persistent net credit or debit balances.

In the usual BFH scenario, it is assumed that banks would agree to redeem their checks, deposits, notes, or coins with gold equal in value to the bundle of goods that defines the dollar. By the same principle, a bank could insist on collecting on a net credit balance at the clearinghouse by obtaining gold equal in value to the bundle.

Index future convertibility would require that banks accept a slightly different obligation. To make the payments system as consistent as possible with Sumner's proposal (index future convertibility with parallel open market operations), suppose banks agreed to redeem checks, deposits, notes, or coins with T-bills of equal market value and an index futures contract on the CPI. A member of the public could use validly drawn checks, notes, or coins to obtain T-bills and a long position on the CPI futures contract at the issuing bank. The bank would be obligated to provide the T-bills and take the matching short position.

"Deposits" are less obvious, but a plausible rule would have the deposit of T-bills and a short position at a bank imply that a customer has sufficient funds to write a check or withdraw notes and coins of equal value. The bank would be obligated to accept the T-bills and take the matching long position.

As with other versions of BFH, a similar obligation would apply at the clearinghouse. A bank with a net credit balance could insist it be settled with T-bills and a long position on the CPI contract.⁴

³A rationale for banks accepting each others' notes for deposit was provided by Lawrence H. White (1984, pp. 18-21).

⁴As explained above, it could instead do nothing or use conventional liquidity management to offset the balance by adjusting its asset or liability portfolio as it prefers. Continuous clearing with treasury bills and index futures (or gold in the usual BFH scenario) would be inconvenient, since it would require constant monitoring of the price of the settlement medium and some extra calculation cost. (White 1986 emphasized similar difficulties.)

The T-bills would be provided by the banks with the matching net debit balances and they must also take the opposite short positions, presumably in proportion to the sizes of their balances. Similarly, a bank with a net debit balance could insist on settling the balance with T-bills and a short position. The T-bills and matching long positions would be distributed to the banks with net credit balances.

To sum up, this version of BFH would require that banks offering dollar-denominated checkable deposits, notes, and coins provide two-way convertibility with T-bills and the futures contract. Problems due to the information lag would be avoided by having redemptions and deposits during the current month use the futures contract for the next month's CPI. The number of contracts to be transferred would be calculated by dividing the value of the T-bills by the defined value of the index.

For example, the public could redeem checks, notes, or coins in May by using them to purchase T-bills and CPI futures contracts from the issuing bank. Suppose the market price of a T-bill maturing in one year was five percent, the futures contract was defined to be \$50 times the CPI, and the target for the CPI was 100. The bank would go short on 1.9 ($\$9,500/\$5,000$) June CPI contracts for every \$10,000 face value of T-bills it was required to sell. The CPI would be measured in June and announced in July. If the CPI were 105, the bank would pay \$250 ($5 \times \$50$) per contract. If the CPI were 98, the member of the public would pay the bank \$100 ($2 \times \$50$) per contract.

Stabilizing Speculation

Index future convertibility implies a market process that would keep the expected value of the CPI on target. Suppose a speculator expected inflation. He would withdraw T-bills from his bank and obtain a long position on the futures contract, anticipating a profit on the futures contract. Assuming risk aversion, his bank would attempt to offset its matching short position, even if it had no expectation that the CPI would deviate from target. It would sell bonds, obtain net repayments of loans, and increase its deposit interest rate to obtain a net credit balance in its clearing account. The speculator's bank would then demand the balance be settled with T-bills and a long position on the CPI contract, effectively shifting its short position to the other banks. Assuming the other banks were also risk-averse, they would attempt to obtain net credit balances to allow them to also offset their short positions.

The withdrawal of T-bills by the speculator and the contraction of credit by the banks would decrease the quantity of money. The increase in deposit interest rates would increase the demand for money. To build money holdings, the public would sell other assets and restrict expenditures on goods and services. The resulting downward pressure on prices would reduce the likely value of the CPI at its next measurement. The contraction of money and credit, the increase in interest rates, and the restraint on spending would continue until some bank or member of the public accepted a short position to match the speculator's long position on the futures contract.

If a single speculator expected deflation, he would deposit T-bills at his bank and obtain a short position on the futures contract. Risk aversion implies his bank would seek a net debit balance by buying bonds, expanding loans, and lowering its deposit interest rates in order to shift its long position to other banks. But the other banks would do the same. The expansion in money and credit, the lower interest rates, and the stimulus to spending would continue until some bank or member of the public accepted a matching long position.

In equilibrium, everyone might agree that credit conditions, the quantity of money, and interest rates are such that the CPI could most probably be on target at its next measurement. The more likely scenario, however, would be for the amount of funds risked by those expecting the CPI to be above target to just match the amount risked by those expecting the CPI to be below target.

Anticipation of the Consequences of Convertibility

The existence of a market process by which index future convertibility would cause the expected value of the CPI to remain on target would be essential. Anticipation of the process, however, would make redemptions and positions on the futures contract unnecessary.

Suppose the CPI was expected to be above target. Speculation using the futures contract would lead to a contraction of credit and an increase in the interest rate. That would imply capital losses on bond portfolios in the near future, so banks and the public would have an incentive to immediately sell bonds on the market in anticipation. The interest rate on bonds would increase immediately, tending to pull up rates on loans and deposits as well. That would depress spending and prices, tending to keep the CPI on target.

Similarly, if CPI was expected to be below target, speculators would buy bonds to obtain the capital gains that would result from

speculation using index futures. As a consequence, however, interest rates would decrease immediately. Spending would increase, again, tending to keep the CPI on target.

Speculation on the bond market would determine interest rates on bonds (and hence on loans and deposits) such that the average expectation would be for the CPI to remain on target. Needed changes in credit, money, and spending could occur without redemptions of T-bills or positions being taken on the futures contract.

Benefits of Combining Index Futures with BFH

Combining index future convertibility with the BFH would free the financial system from dependence on the Fed's issue of base money. That freedom would provide important benefits, even compared to Sumner's proposal that the Fed utilize index future convertibility.

Since keeping the expected value of the CPI on target usually would require a growing quantity of base money, Sumner's scheme usually would require the Fed to make open market purchases. If the Fed's open market purchases were literally parallel, matching its purchases of futures contracts dollar-for-dollar, then it would be left with a long position matching the needed monthly increase in the quantity of base money. Sumner, however, proposed limiting the Fed's and the speculators' risk by giving the Fed discretion to make open market operations in whatever amount it believed necessary to offset its own long or short position.

Rather than making speculators take some position on the futures contract *first* and only then allowing the Fed to respond with needed open market operations, a more reasonable approach would give the Fed further discretion to *initiate* open market operations. If the Fed created a quantity of base money such that the expected value of the CPI remained on target, speculators would have no reason to take positions on the futures contract. If the Fed failed to effectively target the CPI, however, speculators could take long or short positions on the futures contract. The Fed would then be required to use open market operations to offset its short or long position.

The potential benefit of Sumner's proposal, therefore, would be to create corrective market forces that would limit the mistakes of the Fed. Unfortunately, speculators could not profit from correcting the mistakes of the Fed, because it would free ride on their average expectation by adjusting the quantity of base money in a way that made their average expectation wrong. That would have a peculiar consequence. Speculators would correct the Fed's errors only if at

least one speculator anticipated that at least one other speculator underestimated the size of the needed correction. Otherwise, no speculator would have an incentive to take the position on the futures contract needed to force the Fed to make corrective open market operations.⁵ Furthermore, the Fed's free riding would prevent speculators from investing sufficient resources in gathering and interpreting information to effectively target the CPI.

The BFH system of free banking would avoid those difficulties because speculators would cause appropriate changes in money demand and supply, interest rates, spending, and the expected value of the CPI by trading bonds. Successful speculators would avoid capital losses and earn capital gains on their bond portfolios, even if their discretionary activity caused the CPI to remain on target so that no one could profit from long or short positions on the index futures contracts. Since there would be no base money that could change only with discretionary actions by the Fed or in response to transactions with futures contracts, BFH would harness market forces more directly to targeting the CPI.

Sumner explains that index future convertibility could be used to stabilize any of the economic aggregates that have been proposed by modern macroeconomists in place of the CPI (1991, pp. 754–57). Were BFH combined with index future convertibility, free banking also could be combined with targets for nominal income, a wage index, or some function of unemployment and inflation.

Conclusion

Sumner claims that the information lag creates problems for the BFH system of free banking. He advocates index future convertibility as a solution to the information lag. This comment corrects some errors and misunderstandings in his discussion of BFH and argues that combining index future convertibility with the BFH system of free banking would be a superior approach for providing macroeconomic stability.

References

- Black, Fischer. "Banking in a World Without Money: The Effects of Uncontrolled Banking." *Journal of Bank Research* 1 (Autumn 1970): 9–20.
- Black, Fischer. *Business Cycles and Equilibrium*. New York, N.Y.: Basil Blackwell Press, 1987.

⁵If the speculators must first go short on the contract before the Fed could make any open market purchases, that problem would be even more serious.

- Fama, Eugene F. "Banking in the Theory of Finance." *Journal of Monetary Economics* 6 (January 1980): 39–59.
- Fama, Eugene F. "Financial Intermediation and Price Level Control." *Journal of Monetary Economics* 12 (July 1983): 7–28.
- Glasner, David. *Free Banking and Monetary Reform*. N.Y.: Cambridge University Press, 1989.
- Greenfield, Robert L., and Yeager, Leland B. "A Laissez-Faire Approach to Monetary Stability." *Journal of Money, Credit, and Banking* 15 (August 1983): 302–15.
- Greenfield, Robert L., and Yeager, Leland B. "Competitive Payments Systems: Comment." *American Economic Review* 76 (September 1986): 848–49.
- Greenfield, Robert L., and Yeager, Leland B. "Can Monetary Disequilibrium Be Eliminated." *Cato Journal* 9 (Fall 1989): 405–21.
- Hall, Robert E. "Explorations in the Gold Standard and Related Policies for Stabilizing the Dollar." In *Inflation: Causes and Effects*. Edited by Robert E. Hall. Chicago: University of Chicago Press, 1982.
- Hall, Robert E. "Optimal Fiduciary Monetary Systems." *Journal of Monetary Economics* 12 (July 1983): 33–50.
- Sumner, Scott. "The Development of Aggregate Economic Targeting." *Cato Journal* 10 (Winter 1991): 747–59.
- Sumner, Scott. "Using Future Instrument Prices to Target Nominal Income." *Bulletin of Economic Research* 41 (April 1989): 157–62.
- Sumner, Scott. "The Forerunners of 'New Monetary Economics' Proposals to Stabilize the Unit of Account." *Journal of Money, Credit, and Banking* 22 (February 1990): 109–18.
- Thompson, Earl. "A Perfect Monetary System." Paper presented at the Liberty Fund/Manhattan Institute Conference on Competitive Monetary Regimes, New York, March 1986.
- White, Lawrence H. *Free Banking in Britain*. N.Y.: Cambridge University Press, 1984.
- White, Lawrence H. "Competitive Payments Systems: Reply." *American Economic Review* 76 (September 1986): 850–53.
- Woolsey, W. William, and Yeager, Leland B. "Is There a Paradox of Indirect Convertibility." Paper presented at the annual meetings of the American Economic Association, New Orleans, January 1992.
- Yeager, Leland B. "Deregulation and Monetary Reform." *American Economic Review* 75 (May 1985): 103–7.

INDEX FUTURE CONVERTIBILITY: REPLY TO WOOLSEY

Scott Sumner

W. William Woolsey's critique of my survey of the history of economic aggregate targeting in the *Cato Journal* (Winter 1991) raises a number of interesting issues.¹ I argued that much of the intellectual history of monetary policy rules can be read as an attempt to grapple with the problems raised by information lags in the reporting of economic aggregates. I discussed previous policy proposals from this perspective and ended with a set of proposals by myself, and others,² that would overcome the information lag by targeting the price of futures contracts linked to the relevant economic aggregate.

Woolsey begins by correctly pointing out that I mistakenly described the "BFH" system as having an abstract unit of account and that I also described the system as being a hybrid of proposals by Fisher Black, Eugene Fama, and Robert Hall, whereas it is actually a creation of Robert Greenfield and Leland Yeager (1983).

In my original paper I was unsure as to whether Greenfield and Yeager intended to link their currency to a comprehensive bundle of goods and services, or to a narrower bundle of commodities with contemporaneously available prices. Woolsey argues that the comprehensive bundle was "clearly" what Greenfield and Yeager had in mind. I accept Woolsey's interpretation although I would note that Greenfield and Yeager could have eliminated any ambiguity by

Cato Journal, Vol. 12, No. 2 (Fall 1992). Copyright © Cato Institute. All rights reserved. The author is Associate Professor of Economics at Bentley College.

¹The title of my 1991 paper should have read "The Development of Economic Aggregate Targeting."

²This idea seems to have been independently developed by a number of authors. The proposal in my 1989 paper focuses on overcoming the information lag in the context of a nominal income rule. Glasner (1989) focused on a variant of Fisher's (1920) Compensated Dollar Plan. Glasner credits Earl Thompson for having suggested the idea. Hall's (1983) model stabilizes the price level by allowing expectations to affect money demand, rather than the supply of money. Barro (1979, pp. 30–31) also anticipated certain aspects of the futures targeting approach.

defining the bundle as “comprehensive” rather than “comprehensive enough” for the price level to be “practically steady” (1989, p. 409). Since they did not explicitly analyze the information lag it was easy to read into their proposal a desire to find a commodity bundle with an average price that closely mirrors the overall CPI.

Regarding Woolsey’s criticism of my statement that in the BFH system “adjustments [to gold prices] based on incipient changes in the target price index should, by themselves, be sufficient to maintain price level stability” (1991, p. 752), I would point out that it was Greenfield and Yeager that suggested this mechanism. I merely pointed out that their advocacy of also using “changes in the price of gold itself” as an adjustment mechanism is redundant. It is not a different mechanism, it is the same mechanism viewed from a different perspective.

Woolsey raises a more serious criticism in his denial of my assertion that the BFH system is susceptible to the same sort of destabilizing speculation as would occur under a Fisherian Compensated Dollar Plan. I saw the speculation problem as arising from differences between the official price of gold (linked to the last reported price index) and the market price of gold (affected by expectations of the next price index). Woolsey suggests that this problem would not occur under the BFH system since there would be no official price of gold. He argues that

The information lag could cause a disaster in the BFH system for a different reason. If the price of the bundle actually rose above \$1, the banks would be obligated to sell gold for less than its market price. But for how long? It would be absurd for banks to suffer the consequent losses on all redemptions occurring between the announcement of a high price for the bundle and the subsequent announcement. In the context of monthly announcements, such a requirement would surely result in the collapse of the banking system [Woolsey 1992, 478].

I agree with Woolsey except in his assertion that this process is different from the one that I had described. I would suggest that the price at which “banks would be obligated to sell gold” is the official price of gold.

Woolsey’s Proposal

The bulk of Woolsey’s paper concentrates on combining the idea of an index futures-based monetary system with the BFH system. For readers unfamiliar with this literature it may be helpful to think of Woolsey’s system as one where the medium of account involves a futures contract linked to future announcements of the CPI,

whereas the original BFH system used the basket of goods that comprise the CPI as the medium of account.

In my earlier papers on this subject I assumed the existence of a monetary authority in order to focus attention on the idea of index future targeting. I would agree that combining index future convertibility with a free banking system could well produce greater economic efficiency than would occur under a central banking system. Nevertheless, Woolsey appears to have overstated the advantages of this approach in his concluding paragraphs.

Woolsey (1992, p. 483) argues that under my proposal if “the expected value of the CPI remained on target, speculators would have no reason to take positions on the futures contract.” But surely investors would be expected to have heterogeneous expectations. He goes on to argue that since investors would not expect to profit from these contracts they would not be expected to invest sufficient resources into gathering information relevant to the CPI. The argument that information is a public good has been applied to financial markets in general and would seem equally applicable to Woolsey’s proposal. Woolsey (p. 484) suggests that the fact that under his proposal “successful speculators would . . . earn capital gains” somehow differentiates the proposal from other index futures schemes. But successful speculators would also gain from my scheme. Woolsey needs to explicitly show how a specific profit opportunity, in a specific market, would be available under his proposal, but not under my proposal.

I have one other minor suggestion regarding the Woolsey paper. Woolsey briefly discusses the role of deposits in his system. All that is required to fix the price level, however, is that currency and coins are convertible into index futures.

Extensions of the Basic Model

In my *Cato Journal* paper, I mentioned almost as an aside that the Federal Reserve should conduct parallel operations in Treasury securities. Since Woolsey has shown the importance of this assumption I would like to discuss this issue further.

Assume that the Federal Reserve creates a CPI futures targeting system and that \$300 billion in base money must be created to bring the price level up to its target value. In that case the public would sell \$300 billion worth of CPI futures to the Fed in exchange for \$300 billion worth of cash. Because this cash could be invested in assets earning a positive rate of return, the price of the CPI futures

would be less than the expected future value of the price level.³ This problem could be minimized by simply charging the public a small margin requirement (perhaps T-bills) on CPI futures transactions.⁴ No cash would be exchanged until the settlement date. This, however, would break the link between futures transactions on the money supply.

Combining parallel operations in treasury securities with the futures targeting system would minimize the discrepancy between the futures price and the future expected price level, but still allow futures transactions to have the desired impact on the money supply. In this system the Federal Reserve would passively buy and sell CPI futures at the target price. It would then simultaneously buy and sell treasury securities with the goal of preventing the public (and Federal Reserve) from having a substantial net long or short position in the CPI futures market. In the previous example, the purchase of \$300 billion in treasury securities would accomplish that goal.

Both Woolsey and I missed the implication of index futures convertibility for the time inconsistency problem. In contrast to our current system, a system incorporating index futures would enable investors to profit from a (correct) prediction that the Federal Reserve would miss its announced policy target. This would clearly increase policy credibility *for the period being targeted*. It is not clear that it would eliminate the time inconsistency problem, however, since the expected price level several months into the future could still exceed the target value if the Federal Reserve was expected to abandon the index futures system at some time in the future.

Even the long-term time inconsistency problem could be eliminated if the Federal Reserve offered to buy and sell unlimited quantities of CPI futures for all future months, at the target price. This would not violate the dictum of one target per monetary policy instrument since the money supply would only be adjusted on the basis of purchases and sales of the next month's contract. The public's position on the long-term contracts would, however, constrain the Federal Reserve to adhere to its announced CPI target. For instance, suppose the public expected the Federal Reserve to adhere to the target for a few months and then later abandon CPI futures in order to inflate the currency. The public would then buy long-term CPI

³Kevin Dowd (1992) pointed out this problem. His paper provides an alternative solution.

⁴This sort of margin requirement would also be more consistent with the operation of actual futures markets.

futures. Although this would not directly affect the money supply, these purchases would impose a large potential capital loss on the government. The purchases would continue as long as the perceived gain to policymakers from inflating exceeded the capital losses that would result from abandoning the CPI target.⁵

Brian Hillier (1989) showed that the time-inconsistency problem can result from policymakers having too few policy instruments to achieve the first best solution. Because long-term forward CPI targeting can be separated from one period forward targeting that is utilized solely for money supply control purposes, this would provide the monetary authority with the additional policy instrument necessary to overcome the time inconsistency problem. It would be interesting to see how Woolsey's plan could be amended to deal with this issue.

Conclusion

One clear implication of the index future convertibility literature is that *all* forms of discretionary monetary policy, intermediate targeting, and monetary feedback rules are inefficient. Because of my previous work in this area I am obviously pleased that Woolsey has endorsed the viability, and efficiency, of index future convertibility. In another sense, however, I am disappointed. Despite the fact that (as far as I am aware) no one has challenged the optimality of this system, the profession continues to advocate a wide variety of policy proposals that would appear to be dominated by index future convertibility.⁶ I would welcome the opportunity to respond to a cogent critique of this system.

Woolsey's paper represents a useful addition to the index futures convertibility literature. I appreciate his corrections of several errors in my earlier paper, although, as noted previously, I do not agree with all of his criticisms. In my view, monetary policy research should now concentrate on identifying the proper goals of monetary policy. This will necessarily require a resolution of the age-old dispute between proponents of new classical and sticky-price models.

References

- Barro, Robert J. "Money and The Price Level under the Gold Standard." *Economic Journal* 89 (March 1979): 13–33.
- Black, Fisher. "A Gold Standard with Double Feed-back and Near Zero Reserves." 1981. Reprinted in *Business Cycles and Equilibrium*. New York: Basil Blackwell Press, 1987.

⁵Sumner (1994) examines the time-inconsistency problem in more detail.

⁶McCulloch (1991) provides a good recent example of a monetary feedback rule.

- Dowd, Kevin. "A Proposal to End Inflation." Unpublished Manuscript, University of Nottingham, 1992.
- Fama, Eugene. "Financial Intermediation and Price Level Control." *Journal of Monetary Economics* 12 (July 1983): 7–28.
- Fisher, Irving. *Stabilizing the Dollar: A Plan to Stabilize the General Price Level without Fixing Individual Prices*. New York: Macmillan, 1920.
- Glasner, David. *Free Banking and Monetary Reform*. Cambridge: Cambridge University Press, 1989.
- Greenfield, Robert L., and Yeager, Leland B. "A Laissez-Faire Approach to Monetary Stability." *Journal of Money, Credit and Banking* 15 (August 1983): 302–15.
- Greenfield, Robert L., and Yeager, Leland B. "Can Monetary Disequilibrium Be Eliminated?" *Cato Journal* 9 (Fall 1989): 405–21.
- Hall, Robert. "Optimal Fiduciary Monetary Systems." *Journal of Monetary Economics* 12 (1983): 33–50.
- Hillier, Brian. "Time Inconsistency and the Theory of Second Best." *Scottish Journal of Political Economy* 36 (August 1989): 253–65.
- McCulloch, J. Huston. "An Error Correction Mechanism for Long-Run Price Stability." *Journal of Money, Credit and Banking* 23 (August 1991): 619–24.
- Sumner, Scott. "Using Futures Instrument Prices to Target Nominal Income." *Bulletin of Economic Research* 41 (1989): 157–62.
- Sumner, Scott. "The Development of Aggregate Economic Targeting." *Cato Journal* 10 (Winter 1991): 747–59.
- Sumner, Scott. "The Impact of Futures Instrument Price Targeting for the Precision and Credibility of Monetary Policy." *Journal of Money, Credit and Banking* (1994): forthcoming.
- Woolsey, W. William. "The Search for Macroeconomic Stability: Comment on Sumner." *Cato Journal* 12 (Fall 1992): 475–85.

A TOUGH PUMP TO PRIME

Peter A. Schulkin

As the Federal Reserve guided interest rates lower during 1991, many Fed officials were surprised that their primary measure of the money supply, M2, did not expand faster than it did (e.g., Wessel 1991). The 1991 increase in M2 was 2.8 percent (see Table 1)—near the low end of the Fed's 1991 target range of 2.5 percent to 6.5 percent.

The slow growth of M2 during 1991 stemmed in part from the fact that the Fed controlled and continues to control only a small segment of M2. As a result, economic forces outside of the Fed's control can have significant and unpredictable effects on M2. During 1991, the Fed caused the part of M2 under its direct control to grow by 8.0 percent. At the same time the part of M2 not directly controlled by the Fed grew by only 1.7 percent, frustrating Fed efforts to reach at least the mid-point of its M2 target range.

Segment of M2 Controlled by the Fed

The Fed principally exercises its influence on interest rates and the money supply by making open market purchases or sales of Treasury securities. In this way the Fed can increase or decrease the reserves of the banking system. Under the current schedule of reserve requirements, changes in reserves lead to approximately proportional changes in the total of those deposits with Fed reserve requirements, assuming excess reserves remain stable.

Today, depository institutions have no reason to build excess reserves above negligible amounts since no interest is paid on

Cato Journal, Vol.12, No. 2 (Fall 1992). Copyright © Cato Institute. All rights reserved.

The author is Lecturer in Economics at the University of California, Los Angeles.

TABLE I
CHANGES IN M2 AND ITS PRINCIPAL COMPONENTS 1990 IV TO 1991 IV
(BILLIONS OF DOLLARS, SEASONALLY ADJUSTED)

	M2	Currency in Circulation	Demand Deposits	Other Checkable Deposits	Zero Reserve Requirement Deposits ^a	Non-Institutional Money Market Mutual Funds	Other ^b
1990 IV	3,335.7	245.5	277.5	292.7	2,087.6	346.2	86.2
1991 IV	3,429.3	266.0	287.0	329.1	2,107.9	359.8	79.5
Percentage Change	2.8	8.4	3.4	12.4	1.0	3.9	-7.7
Percent of M2 1991 IV	100.0	7.8	8.4	9.6	61.5	10.5	2.3

^aSavings accounts, money market deposit accounts, and small denomination (less than \$100,000) time deposits.

^bTravelers checks of nonbank issuers, overnight Eurodollar borrowings from U.S. residents by foreign branches of U.S. banks, and overnight repurchase agreements of depository institutions.

SOURCE: Author's calculations using data from Board of Governors of the Federal Reserve System, *Federal Reserve Statistical Release H.6* (26 March 1992).

reserves¹ and since there are no liquidity reasons for holding excess reserves. (Liquidity requirements can be met, for example, by interest-bearing marketable securities and potential short-term borrowings from various sources including the Federal Funds market.) As a consequence, the total amount of excess reserves of all depository institutions has been relatively stable in recent years at very low levels in comparison with total reserves.

Since December 1990, only two types of deposits have been subject to reserve requirements. These are demand deposits and "other checkable deposits," both of which were subject to the same reserve requirements during 1991—3 percent of the first \$41.1 million of such accounts at a depository institution and 12 percent over \$41.1 million.²

Despite the fact that the payment of interest on demand deposits is prohibited, they are still a significant component of M2 at 8.4 percent of the total for 1991 IV. The results of a Federal Reserve study (Mahoney 1988) indicate that most of these zero interest deposits are held by businesses and count as compensating balances. Part of the business usage of demand deposits is attributable to regulations which restrict corporate checking accounts to demand deposits. However, the payment of fees has been gradually replacing the use of compensating balances, especially for the large customers of large banks, some of whom have their demand deposit balances automatically swept out at the end of the day to earn interest overnight. As a result, the increase in demand deposits outstanding in recent years has been limited. Even after a 3.4 percent growth in 1991, demand deposits outstanding at year-end were at a lower level than they were five years earlier. Due to the factors restricting the growth of demand deposits, the impact of Fed open market operations during 1991 was much greater on the only other deposit category with reserve requirements—the other checkable deposits category, which

¹If a rate of interest close to short-term market rates is paid on reserves, excess reserves would likely become much more significant and volatile than they are today. This would probably make the Fed's job of managing a stable money supply more difficult. Nevertheless, as discussed in "Greenspan Asks that Fed Be Allowed to Pay Interest" (1992), Federal Reserve Chairman Alan Greenspan sought Congressional approval for the Fed to pay interest on reserves, arguing that zero interest on reserves amounts to a tax passed on to bank depositors and borrowers encouraging both to bypass the banking system.

²These were the reserve requirements in effect during most of 1991. Effective in late December 1991 the break point for shifting from 3 percent to 12 percent was increased to \$42.2 million and effective April 1992 the maximum rate was decreased from 12 percent to 10 percent.

increased by 12.4 percent during 1991 ending the year at 9.6 percent of M2.

The small part (18 percent for 1991 IV) of M2 represented by these two deposit categories with reserve requirements is the most responsive to Fed open market operations. Under the present system of reserve requirements this response is relatively predictable—the total of deposits with reserve requirements will change by approximately the same percentage as the percentage change in reserves. During 1991 the reserves of depository institutions increased by about 8.9 percent,³ and for the same period the total of deposits with reserve requirements (demand deposits plus other checkable deposits) increased by 8.0 percent.

Segments of M2 outside of Fed Control

For the purposes of this exposition those segments of M2 outside of the Fed's control are divided into three categories—currency in circulation, zero reserve requirement deposits (ZRRDs), and non-institutional money market mutual funds.

Currency in Circulation

One of three main categories of M2 outside of the Fed's control is currency in circulation. The Fed passively supplies currency to or accepts currency from depository institutions, balancing by debiting or crediting depository institution accounts at the Fed. The amount of currency in circulation increased by 8.4 percent in 1991, and it ended the year at 7.8 percent of M2.⁴

Zero Reserve Requirement Deposits Included in M2

The second and largest category of M2 outside of the Fed's control consists of zero reserve requirement deposits. For 1991 IV the ZRRDs of banks and other depository institutions (hereafter referred to as "banks") accounted for 61.5 percent of M2. Included in this category are three types of deposits: money market deposit accounts, savings accounts, and small denomination (less than \$100,000) time

³Based on data adjusted for changes in reserve requirements from Board of Governors (1991) and (1992). In general, the percentage changes for reserves and deposits with reserve requirements are not equal due to slippages related to how the data is compiled and changes in excess reserves.

⁴If one is relating M1 or M2 to economic activity in the United States, it can be argued that M1 and M2 are overstated to the extent that currency in circulation includes U.S. currency circulating in foreign countries. This overstatement may be significant as it has been estimated that more dollar currency circulates outside the United States than inside the United States (Clark and Malabre 1990).

deposits. Large denomination (\$100,000 or more) time deposits also have zero reserve requirements but are not included in M2.

Since there are no reserve requirements on ZRRDs, the Fed cannot influence their magnitude directly by increasing or decreasing reserves. The amount of ZRRDs outstanding and the rate of interest paid on them will depend on supply and demand factors including alternative rates of return available to economic units with funds to invest (affects the demand for deposits) and the availability of loans and investments that meet the banks' current criteria (affects the supply of deposits).

Widely different factors can affect the supply and demand for ZRRDs. For example, when the Resolution Trust Corporation (RTC) takes over the assets of failed thrifts that are liquidated, and the depositors are paid off, the supply of deposits (primarily ZRRDs) is decreased and M2 will be negatively impacted. Many former depositors of liquidated thrifts will shift their money to other depository institutions, but some will not due to the lower average interest rates paid by the remaining depositories. In addition, the closing of high-rate, insolvent thrifts and the movement of deposits to solvent depository institutions will put a downward pressure on deposit rates, causing some depositors of solvent institutions to shift funds to alternative assets. In the end, the liquidation of failed thrifts will lead to some resultant increase in deposits and assets at the surviving depository institutions, but such increases will not fully offset the losses of deposits and assets of the liquidated thrifts.

If banks are not seeing many loans that meet their lending criteria, they will not aggressively seek ZRRDs, thereby making the supply of ZRRDs lower than it would otherwise be.⁴ Judging from anecdotal evidence, this may have been a significant factor in the 1991 slow growth of ZRRDs.

Increases in the supply of competing assets may also reduce the demand for ZRRDs and, other things equal, the amount of M2. Increases in government debt, corporate debt, and newly issued equity securities all compete for money which might otherwise be invested in ZRRDs. During 1991 there were significant increases in outstanding amounts of all three of the aforementioned competing assets. Moreover, in many instances funds raised from the sale of newly issued corporate debt and equity securities were used (at least initially) to pay down bank borrowings.

Easy money policies designed to *expand* M2 and lower interest rates may put *contractionary* pressures on ZRRDs as depositors take money out of deposits with falling interest rates and buy other assets with potentially higher yields, such as stocks and long-term bonds.

However, some of this asset shift effect is likely to be offset by additions to ZRRDs by the sellers of existing stocks and bonds. In addition, the expected returns of competing assets such as stocks and bonds are likely to decline as interest rates on ZRRDs decline, reducing incentives to switch out of ZRRDs. The ultimate effect of easy or tight money policies on ZRRDs is uncertain.

There is a potential upper bound to the amount of funds that banks can raise from ZRRDs and all other deposits—the bank capital constraint. Losses in asset value or growth in asset holdings will usually result in a decline in capital to assets ratios unless a sufficient amount of new capital is obtained.⁵ Whenever a bank's capital ratios decline to near the regulatory minimums or lower, its growth will be restricted until it obtains more capital.

During 1991, much bank capital was lost due to problem loans, and many banks limited asset growth and even contracted in size in order to meet minimum capital requirements. In extreme cases banks lost all of their capital and failed. In addition, banks were required to meet the two-tiered Basle Accord capital requirements in full by the end of 1992. The impact of these individual bank capital constraints on aggregate ZRRDs is unclear, but there is some evidence that the shortage of bank capital was not a nationwide problem. In those regions where a shortage of bank capital might have limited bank deposits and lending, the effect was at most a modest one (Bernanke and Lown 1991). To the extent that new sound loan opportunities were available, they may have been taken up by banks with capital ratios above the minimum or by banks with access to new capital on acceptable terms.

Capital availability or constraints and factors affecting the supply and demand for ZRRDs can all cause ZRRDs to increase or decrease for reasons outside of the Fed's control. The balance of all of these factors in 1991 led to a very small increase—about 1 percent—in ZRRDs included in M2, which made it more difficult for the Fed to achieve its M2 objectives.

Non-Institutional Money Market Mutual Funds

The third principal category of M2 outside of the Fed's control consists of non-institutional money market mutual funds, which constituted 10.5 percent of M2 as of the last quarter of 1991. Although

⁵The regulatory capital requirements are based on capital to risk-weighted assets ratios. The addition of very low risk assets, such as securities guaranteed by the U.S. government or its agencies, will not reduce a bank's capital ratios since the risk weight assigned to these assets is zero. For information on the risk weights and other aspects of the Basle Accord, see Board of Governors (1989).

these are not bank deposits, holders of these funds can usually write checks (often with some restrictions) against their money invested. For the year 1991, the dollars invested in non-institutional money market mutual funds increased by 3.9 percent, despite the significant decline in market interest rates. The amount of money invested in money market mutual funds is outside of the direct control of the Fed and depends on many factors including the rates of return available on competing assets.

Implications for Monetary Policy

For 1991 IV, about 80 percent of M2 consisted of bank deposits and 86 percent of bank deposits were included in M2. Despite its decreasing role in the economy in recent years, the banking sector remains a significant supplier of credit to the economy and bank deposits continue to constitute the major part of M2. When the Fed focuses on M2 it is in effect primarily focusing on bank deposits or the principal resources of the banking sector.

Considering only bank deposits,⁶ the 1991 money and credit picture becomes bleaker than that portrayed by M2. While M2 increased by 2.8 percent from 1990 IV to 1991 IV, total bank deposits increased by only 0.2 percent over the same period. Without assigning cause and effect one may conclude that the banking sector did not respond positively to the easier money policy of 1991.

The Fed increased total bank reserves in 1991 in an effort to increase the money supply and stimulate the economy. That small part of M2 tied to the level of total bank reserves through reserve requirements increased in step with the increase in reserves. However, the larger part of M2 with no direct tie to total bank reserves did not follow the lead of the Fed. The sluggish response of the banking system to the Fed easing limited the growth of bank lending, bank deposits, and M2. Since monetary policy normally works in part through changes in bank deposits and lending, the fact that it did not do so during 1991 suggests that larger than usual (for this stage in the business cycle) percentage increases in total bank reserves were required to achieve the Fed's objectives.

⁶Several modifications are required to transform M2 into a 100 percent measure of bank deposits in the United States. Currency in circulation would have to be deleted (see footnote 4). Non-institutional money market mutual funds would also have to be deleted but this would be offset in part by a third modification—the addition to M2 of large (\$100,000 and over) time deposits (which are held in significant amounts by money market mutual funds as a group). Three other changes of smaller magnitude would also be necessary—the deletion of overnight Eurodollar deposits of U.S. residents in foreign branches of U.S. banks, the deletion of overnight repurchase agreements of depository institutions and the deletion of travelers checks of nonbank issuers.

References

- Bernanke, Ben S., and Lown, Cara S. "The Credit Crunch." *Brookings Papers on Economic Activity* 2 (1991): 205–47.
- Board of Governors of the Federal Reserve System. *Capital Adequacy Guidelines*, 15 March 1989.
- Board of Governors of the Federal Reserve System. *Federal Reserve Bulletin* 77 (July 1991).
- Board of Governors of the Federal Reserve System. *Federal Reserve Bulletin* 78 (April 1992).
- Clark, Lindley H. Jr., and Malabre, Alfred L. Jr. "Defying Predictions, Supply of U.S. Cash Is Surging." *Wall Street Journal*, 4 October 1990, p. A2.
- "Greenspan Asks That Fed Be Allowed to Pay Interest." *Wall Street Journal*, 11 March 1992, p. C9.
- Mahoney, Patrick. "The Recent Behavior of Demand Deposits." *Federal Reserve Bulletin* 74 (April 1988): 195–208.
- Wessel, David. "Money-Supply Puzzle Is Troubling the Fed." *Wall Street Journal*, 7 October 1991, p. A1.
- Wenninger, John, and Partlan, John. "Small Time Deposits and the Recent Weakness in M2." Federal Reserve Bank of New York. *Quarterly Review* 17 (Spring 1992): 21–35.

BOOK REVIEWS

Generational Accounting: Knowing Who Pays, and When, for What We Spend
Laurence J. Kotlikoff
New York: Free Press, 1992, 220 pp.

Budget wonks may be the only people to have noticed a new chapter in the FY 1993 federal budget, one titled "Generational Accounts Presentation." Seldom has a new economic idea moved so fast from conception to semi-official status, so it is important for the broader community to understand what is at issue.

Generational accounting is a framework for estimating the distribution of payments and receipts across generations from any specific set of government fiscal policies. Laurence Kotlikoff, the primary developer of this concept, has written this book for the intelligent lay reader. He explains why the current budget concepts are inadequate or misleading and lays out the implications of this new concept. His book is largely successful on both counts.

Conventional budget accounting is on a cash-flow basis. The purchase of new assets is treated as an expense, and existing assets are not depreciated. Current commitments for future expenses, such as for government employee pensions and loan and deposit guarantees are not treated as an expense until the payments are made. This budget concept is useful for cash management, but it seriously misrepresents the level and time distribution of the cost of current government services and it invites the abuse of financing current services from future payments.

The national income concept of the budget, in contrast, is on a partial accrual basis. Current expenditures are reduced by the amount of net loans and the purchase of other financial assets and are increased by the amount of commitments for government employee pensions. In recent years, for example, the deficit on the national income basis has been substantially smaller than on the budget basis, primarily because of the large amount of payments to purchase the assets of failed S&L associations.

For many years, based on the theory developed by John Maynard Keynes in the 1930s, current government expenditures were considered an addition to total demand in the economy, and current tax receipts were considered as a reduction to total demand. The deficit, thus, was considered the best measure of net fiscal stimulus and, depending on the state of the economy and financial markets, was expected to increase

some combination of output, employment, the general price level, and interest rates.

Given this perspective, which of the two budget concepts above is the best measure of the economic impact of the government budget? The answer, maybe surprisingly, is neither! When I first made such comments at a public meeting in 1981, it provoked a storm of protest similar to that provoked by the boy who observed that the emperor has no clothes, and I was nearly fired from my position as a White House economic adviser. Since that time, there has been a large empirical literature to test the economic effects of fiscal policy and the development of a new budget concept by Robert Eisner in an attempt to rescue Keynesian macroeconomics.

Kotlikoff's response to this debate has been to start all over, to develop a framework for evaluating the economic effects of fiscal policy on a microeconomic foundation. He pulls no punches; the first sentence of this book asserts that "the government's budget deficit . . . is a number devoid of economic content and that its use has repeatedly led us astray." He is indifferent to the measured deficit but not to current fiscal policy, because he demonstrates that future generations will have to pay at least 21 percent more of their lifetime incomes than today's young workers in order to maintain current policies. Moreover, he demonstrates that most of this burden on future generations is a consequence of policies approved between 1950 and 1979, a period in which the ratio of the public debt to current output declined almost continuously.

The first part of this book is a devastating critique of Keynesian macroeconomics and of the leading apocalypics among contemporary economists that will either delight or offend the reader. Kotlikoff then argues that each of the several definitions of the deficit is arbitrary, and he summarizes a large body of empirical tests of the Keynesian conjectures. My sense is that he has overstated his case on these issues. The several definitions of the deficit are arbitrary for some purposes but not for others. Kotlikoff summarizes the growing body of studies that are inconsistent with the Keynesian conjectures but does not mention or evaluate any of the confirming evidence. Some simple studies of my own, for example, find a significant positive effect of (the national income measure of) the deficit on the real corporate bond rate and on the U.S. net foreign balance. A broader survey of the empirical studies provides only mixed evidence for Kotlikoff's unqualified conclusions; there is still a basis for a legitimate dispute among economists on these issues.

Whatever one's views about the economic effects of the measured deficit, Kotlikoff makes an effective case that his framework of generational accounting provides better estimates of the intergenerational effects of the full range of fiscal policies. He is also careful to caution readers about drawing normative conclusions from his results. He first captures the reader's attention by summarizing the generational accounts for the policies in effect in 1989, results that should shock anyone under

the age of 40. He then demonstrates the power of this technique by simulating the effects of five specific policy changes and the combination of policy changes in each of the past four decades.

These results are very interesting by themselves, rather like reading the national income accounts for the first time. The appropriate next stage of this research is to test whether these new accounts help explain the fiscal effects on a range of current economic conditions. For the moment, Kotlikoff has left this task for another day or for other scholars.

William A. Niskanen
Cato Institute

The Development Frontier: Essays in Applied Economics

Peter Bauer

Cambridge, Mass.: Harvard University Press, 1991, 241 pp.

Peter Bauer has, for a half-century, employed a sharply critical intellect, fortified by training in economic reasoning, examination of the literature, and field researches done most intensively in Southeast Asia and West Africa, to challenge stereotypical ideas in the popular, political, and academic minds about the causes of poverty in the world.

His writings have had substantial influence. The notion that the improvement of the conditions of life of the poor requires the aggressive intervention of the state in markets has now, because of his writings, much less currency and acceptance than it did some decades ago. Since the state's intervention often had adverse consequences, he has rendered a signal service to the poor of the world.

The politico-economic environment within which he has written his papers may, among others, be characterized by following features: (1) Government policy in the poor countries has been interventionist. This has distorted the composition of economic output, diminished economic growth rates, and intensified poverty. The policy has reflected a mistrust of markets where exchanges are consensually consummated, the terms of exchange are set by market forces, and there is freedom to exit and enter. (2) Interventionist policy has been generated by a number of variables, such as the state's legal monopoly in the exercise of coercive power and the officeholders' penchant for imposing behavioral rules. (3) Public officeholders perceive their role as benign. They fail to see that their behavior can have unforeseen, adverse, and malign consequences. (4) Rule-making is often a source of income for public officeholders. Those in the community who would be privileged by a rule may be willing to pay those who exercise rule-making authority. By lobbying legislatures and officeholders, special interests secure governmental arrangements that serve them at the expense of others. (5) In most of the poor countries that were once colonies, the traditions of

control and economic intervention established by the ruling metropolitan powers survived the end of colonial rule. (6) Eastern European countries were governed aggressively by the Soviet Union and its proxies in their own countries, and an elaborate ideology of interventionist state behavior was drummed into the psyche of the population. These features served (7) to develop a set of flawed ideas that have been promoted by the intellectual communities of the poor countries and by the community of international civil servants. Those ideas have been seized by some sectors of the community to tilt policy design toward a system that produces economic rents for them at the expense of other sectors. Moreover, the academic and intellectual communities of those countries have come to be committed to central planning and versions of Fabian socialism as fundamental prescriptive components of policy. Their ideas have had a powerful influence in the design of practical governmental affairs.

The ideas produced by this environment are laid out by Bauer in chapter 17, "Development Economics: A Retrospective View." The paper, initially delivered at a World Bank lecture series, describes the standard view thus: the advancement of less developed countries depends upon supplies of capital; but this capital cannot be generated in the poor countries because incomes are low, virtually the whole of income is consumed, and profitable investment opportunities are sparse; these problems are reinforced by a lack of responsiveness to economic incentives and enterprise. The accepted view is, as Bauer notes in his summary (p.187),

if significant economic advance is to be achieved, governments have an indispensable as well as a comprehensive role in carrying through the critical and large-scale changes necessary to break down the formidable obstacles to growth and to initiate and sustain the growth process. External trade is at best ineffective for the economic advance of the less developed countries and more often it is damaging.

Throughout his long and impressive career Bauer has successfully challenged these stereotypical ideas. He does so again in this collection of seventeen essays. Six of them have been specifically written for this book. Most of the others are revised versions of papers that Bauer had previously published in academic and professional journals.

The essays all deal with experience and policy in poor countries. Within that rubric, they cover a range of interesting topics: among them, the contribution of trading to economic growth, intervention in marketing of primary commodities, price and wage controls, and price and wage effects. The essays are a balanced combination of theory and exposed experience.

They are a splendid sample of Bauer's academic and intellectual contribution to the field in which he has specialized.

Simon Rottenberg
University of Massachusetts

Constitutional Environments and Economic Growth

Gerald W. Scully

Princeton, N.J.: Princeton University Press, 1992, 248 pp.

Proponents of the free market often claim that increased government reduces growth of GNP, that adherence to the "rule of law" increases growth of GNP, and that reduced regulation of economic life increases economic opportunity. In this important book, Gerald Scully of the University of Texas at Dallas has provided exactly the substantiation we would want for these propositions.

After an overview, Scully discusses the economic theory of growth, following from the work of Robert Solow. Scully makes the important point that in that model and its progeny, the institutional structure of the economy (for example, whether it is a market or a planned economy, a democracy or a dictatorship) is irrelevant. Those models focus on economic inputs, primarily labor and capital. He modifies the theoretical model to allow for institutional factors such as the property rights structure of the economy, and he discusses the role of human capital, technological change, and capital formation within the context of growth models.

The book next summarizes the economic theory of constitutions, based on the work of James Buchanan and Gordon Tullock. In anarchy, resources would be consumed in committing and protecting against theft. To avoid those losses, humans institute governments to establish and protect property rights and to provide public goods, the most important of which is defense. Different constitutions define different structures of rights, which have implications for the level and growth of income. As rights (not wealth) become more equally distributed and better protected from appropriation by individuals or government, incomes increase. Chapter 3 also considers alternative legal systems. Following the work of Friedrich Hayek and Richard Posner, Scully distinguishes between common law and statutory, or code, systems, and argues for the relative efficiency of the former.

Scully's next task is to set forth "a model of the evolution of the constitutional setting." In that model we begin with some constitution. Agents then invest in private activities under the rules defined by the constitution. As investment proceeds, the marginal product of capital invested in private activities falls. It eventually pays for agents to invest in rule change. Rule changes can be efficient, but there are also possibilities for rent seeking and other inefficient rule changes. When sufficient numbers of inefficient rules have been put in place, a constitutional revolution may occur that will eliminate these inefficient rules and allow the process to begin again. Scully maintains that this has been the situation in Russia and the Eastern countries, which began with highly inefficient constitutions (forbidding private property) and hence reached the limits of inefficiency and constitutional revolutions relatively quickly.

This theory has an advantage over others (such as Mancur Olson's theory) because it makes the process of progressive rent seeking endoge-

nous to the model and the result of economic factors, primarily the fall in the marginal return to investment in private activities. However, in my view, this is the most controversial assumption in the model because technological change can eliminate this tendency. Scully also recognizes that this assumption is debatable. Nonetheless, the result—that modern economies seem to exhibit increasing amounts of rent seeking and other inefficient investments—is uncontroversial, and this model is an interesting attempt to explain this observation.

In subsequent chapters, Scully discusses alternative measures of political and civil liberty available in the literature and examines the relation between freedom and type of law. The common law countries are defined as Great Britain and her former colonies (a total of 54 countries). All other noncommunist countries (94) are characterized as civil law countries. Marxist and Muslim legal systems are also considered. The major empirical result is that common law countries are significantly more likely to be free than are any other countries. Marxist and Muslim countries are less likely to be free than non-Marxist and non-Muslim countries.

Actual economic behavior is examined for 115 non-Marxist economies for the period 1960-1980, and Scully here presents the most important conclusions in the book. The capitalist paradigm is contrasted with the paradigm calling for increased government intervention to stimulate economic growth. The results (p. 176) are worth quoting:

On average, politically open societies grew at a compound real per capita rate of 2.5 percent per annum, compared to a 1.4 percent growth rate for politically closed societies. On average, societies that subscribe to the rule of law grew at a 2.8 percent rate, compared to a 1.2 percent rate in societies where state rights take precedence over individual rights. On average, societies that subscribe to private property rights and a market allocation of resources grew at a 2.8 percent rate, compared to a 1.1 percent rate in nations where private property rights are circumscribed and the government intervenes in resource allocation.

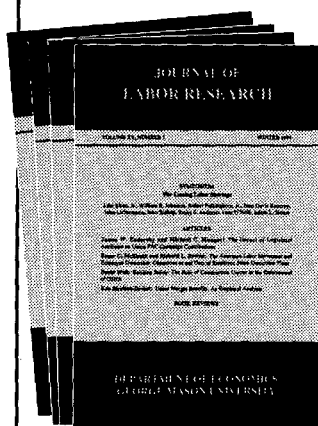
Scully also shows (p. 179) that “societies where freedom is restricted are less than half as efficient in converting resources into gross domestic product as free societies.”

Many who believe that economic freedom leads to higher growth nonetheless favor increased government because they fear that economic freedom leads to excessive concentration of wealth and income inequality. Scully examines that argument and concludes that in fact “income inequality is somewhat more than 50 percent higher in societies where individual rights are restricted” than in freer societies (p. 191-94). He also shows that as the size of government as a percentage of the total economy increases, economic growth decreases significantly, and resources are used less efficiently. Finally, Scully summarizes the policy implications of the analysis and ends on a rather pessimistic note. He concludes that in the United States and Great Britain, the forces of the

1980s for reduced government seem to have run their course, and much of the rest of the world has not adopted these policies.

While this book contains some technical theoretical and statistical analyses, it is useful for others besides professional economists. The conclusions summarized above are sufficiently powerful that anyone interested in the role of freedom in economic development would profit from close attention to this book. Parts are also accessible to undergraduates in courses in economic development, law and economics, government regulation, and in general economic policy courses.

Paul H. Rubin
Emory University



JOURNAL OF LABOR RESEARCH

Selected Articles Forthcoming in Volume XV, 1994

**Cities, Unions, and
the Privatization of
Sanitation Services**

Timothy D. Chandler and
Peter Feuille

**American Unemployment in
Historical Perspective**

Richard K. Vedder and
Lowell E. Galloway

**Unions and Productivity:
Evidence from Academe**

Mark Meador and
Stephen J. K. Walters

**Union Activity and the Decline in American
Trade Union Membership**

C. Timothy Koeller

Deregulation and Labor Earnings

Wallace Hendricks

**American Trade Unions and Public Approval:
Can Unions Please All of the People All of the Time?**

Paul Jarley and
Sarosh Kuruvilla

District Elections and the Power of Municipal Employee Unions

Stephen L. Mehay and
Rodolfo A. Gonzalez

**A Simultaneous-Equations Model of the Relationship
Between Public Sector Bargaining Legislation and
Unionization**

Melissa S. Waters,
R. Carter Hill,
William J. Moore, and
Robert J. Newman

**Quality Upgrading and Restrictive Hiring Practices in
Unionized Jobs**

Jayanthi Krishnan

**Union Premiums in the Federal and Private Sectors:
Alternative Evidence from Job Queues**

Madhu S. Mohanty

Interest Group Tactics and Public Sector Labor Relations

Susan Schwochau

**The Political Effects of Grievance Handling by
Stewards in a Local Union**

David Meyer

**U.S. Unionism and R&D Investment:
Evidence from a Simple Cross-Country Test**

John T. Addison and
Joachim Wagner

Trade Unions and Quits: Australian Evidence

Mark Wooden and
Meredith Baker

**The Political Influence of Unions and Corporations on
COPE Votes in the U.S. Senate**

William J. Moore,
Denise R. Chacere,
Thomas D. Curtis, and
David Gordon

The *Journal of Labor Research* is published quarterly by the Department of Economics of George Mason University. Subscriptions are on a calendar year basis and subscribers will receive all issues for the current year. Annual subscription rates are: individuals, \$35.00; academic libraries, institutions, government agencies, and business firms, \$100.00. Foreign subscribers add \$10.00 to cover additional postage. All remittances should be made payable to *Journal of Labor Research*, Department of Economics, George Mason University, Fairfax, Virginia 22030.

Editorial correspondence and manuscripts should be sent to James T. Bennett, Editor, *Journal of Labor Research*, Department of Economics, George Mason University, Fairfax, VA 22030. Manuscripts should be submitted in triplicate with a 100-word abstract. A submission fee of \$20.00 for nonsubscribers must accompany all manuscripts. Papers accepted for publication must conform to JLR style requirements; style sheet available on request.

Modern scholarly editions of classic works for today's readers

THE LIMITS OF STATE ACTION

NEW By Wilhelm von Humboldt
Edited by J.W. Burrow

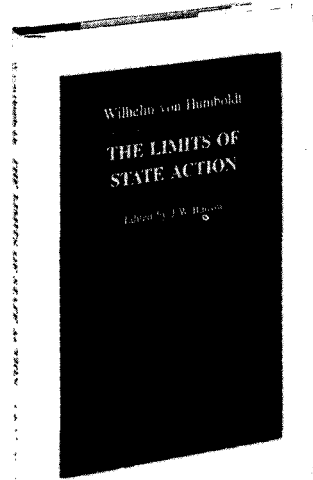
*Nearly two hundred years after its origination, **The Limits of State Action** remains an unsurpassed statement of the case for classical liberalism.*

—Hillel Steiner, University of Manchester

Written in 1792 by “Germany’s greatest philosopher of freedom,” as F. A. Hayek called him, this clear and direct discussion of freedom and individual responsibility argues against state promotion of citizen welfare and state interference with private acts.

162 + lxii pages. Editor’s introduction, select bibliography, index.

Hardcover	\$25.00	0-86597-108-0
Paperback	\$ 7.50	0-86597-109-9



ALSO AVAILABLE

THE PRINCIPLES OF ETHICS

In Two Volumes

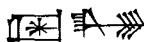
By Herbert Spencer

Introduction by Tibor R. Machan

“The study of Spencer’s ethics can shed needed light on some of the intricacies of what is demonstrably the best perspective on the ethical and political aspects of human life, namely the morality of rational self-interest and the politics of the free society.” —Tibor Machan, in his Introduction

Hardcover	\$35.00 the set	0-913966-33-9
Paperback	\$15.00 the set	0-913966-34-7

To order by phone, call 800-955-8335; fax 317-577-9067. Prepayment by check, VISA, or MasterCard is required on all orders not for resale. We pay book rate postage on orders prepaid by check. Please allow approximately four weeks for delivery.



LIBERTY FUND

8335 Allison Pointe Trail, Dept. Q311, Indianapolis, IN 46250-1687

ELECTRONIC REPRODUCTION PROHIBITED

Journal of Health Politics, Policy and Law

James A. Morone, Editor

The Media and the Fate of the Medicare Catastrophic Extension Act

David P. Fan and Lois Norem

Rust vs. Sullivan in the Supreme Court

Kenneth R. Wing

Why Requiring Employers to Provide Health Insurance Is a Bad Idea

Katherine Swartz

Partnership as a New Strategy in Mental Health Policy: The Case of
Québec

Françoise Boudreau

Negotiating National Health Insurance

Lynn Etheredge

German Unification: East Meets West in the Doctor's Office

Deborah A. Stone

Please enter my subscription (4 issues) to **Journal of Health
Politics, Policy and Law** for 1992. (Add \$8 for subscriptions
outside the U.S.)

_____ Institutions, \$76 _____ Individuals, \$40

_____ Students, \$20 (Please include photocopy of current I.D.)

Payment options:

_____ I enclose my check payable to Duke University Press.

Charge my _____ VISA _____ MasterCard _____ American Express.

No. _____ Expiration _____

Signature _____

Name _____

Address _____

Call or fax us between 8:00 and 4:00 EST with your credit card
order. Phone (919) 684-6837 FAX (919) 684-8644

Send this coupon to: **Duke University Press, Journals Division,**
6697 College Station, Durham, NC 27708

JOURNAL OF MONEY, CREDIT, AND BANKING

Widely read and cited by researchers and policymakers, the *Journal of Money, Credit, and Banking* is a primary economics journal reporting major findings in the study of monetary and fiscal policy, credit markets, money and banking, portfolio management, and related subjects. Editors: Stephen G. Cecchetti and Paul D. Evans, Ohio State University.

Recent and Forthcoming Articles:

Fiscal Policy and the Term Structure of Interest Rates

Walter H. Fisher and Stephen J. Turnovsky

Innovations in Interest Rates, Duration Transformation, and Bank Stock Returns

Srinivas R. Akella and Stuart I. Greenbaum

Debt Management Objectives for a Small Open Economy

Paul Boothe and Bradford Reid

Competitive Externalities and the Optimal Seigniorage Segmentation

Joshua Aizenman

Monetary Aggregates as Monetary Targets

William Roberds and Charles H. Whiteman

Treasury Bill Rates in the 1970s and 1980s

Patric H. Hendershott and Joe Peek

Consumer Durables: Evidence on the Optimality of Doing Nothing

Avner Bar-Ilan and Alan S. Blinder

A Generalized Valuation Model for Fixed-Rate Residential Mortgages

James B. Kau, Donald C. Keenan, Walter J. Muller III, and James F. Epperson

The Currency Substitution Hypothesis and Relative Money Demand in Mexico and Canada

John H. Rogers

Published quarterly by the **Ohio State University Press**

1070 Carmack Road, Columbus, Ohio U.S.A. 43210-1002 614/292-1407

One-year subscription: \$35.00 Individual \$70.00 Library \$20 Student

The Cato Institute

Founded in 1977, the Cato Institute is a public policy research foundation dedicated to broadening the parameters of policy debate to allow consideration of more options that are consistent with the traditional American principles of limited government, individual liberty, and peace. To that end, the Institute strives to achieve greater involvement of the intelligent, concerned lay public in questions of policy and the proper role of government.

The Institute is named for *Cato's Letters*, libertarian pamphlets that were widely read in the American Colonies in the early 18th century and played a major role in laying the philosophical foundation for the American Revolution.

Despite the achievement of the nation's Founders, today virtually no aspect of life is free from government encroachment. A pervasive intolerance for individual rights is shown by government's arbitrary intrusions into private economic transactions and its disregard for civil liberties.

To counter that trend the Cato Institute undertakes an extensive publications program that addresses the complete spectrum of policy issues. Books, monographs, and shorter studies are commissioned to examine the federal budget, Social Security, regulation, military spending, international trade, and myriad other issues. Major policy conferences are held throughout the year, from which papers are published thrice yearly in the *Cato Journal*.

In order to maintain its independence, the Cato Institute accepts no government funding. Contributions are received from foundations, corporations, and individuals, and other

revenue is generated from the sale of publications. The

Institute is a nonprofit, tax-exempt, educational foundation under Section 501(c)3 of the Internal Revenue Code.



The Watterston House,
headquarters of the
Cato Institute

THE DEVELOPMENT FRONTIER

The argument for aid as necessary for development rests on the belief that possession of capital is critical for economic advance. If this were so, how is it that large numbers of very poor people could have become prosperous within a few years without donations, as they have done the world over? Evident examples include very poor immigrant communities in North America and Southeast Asia. To have capital is the result of economic achievement, not its precondition.

—Peter Bauer

*The Development Frontier:
Essays in Applied Economics*
(Cambridge: Harvard University Press,
1991), p. 43.